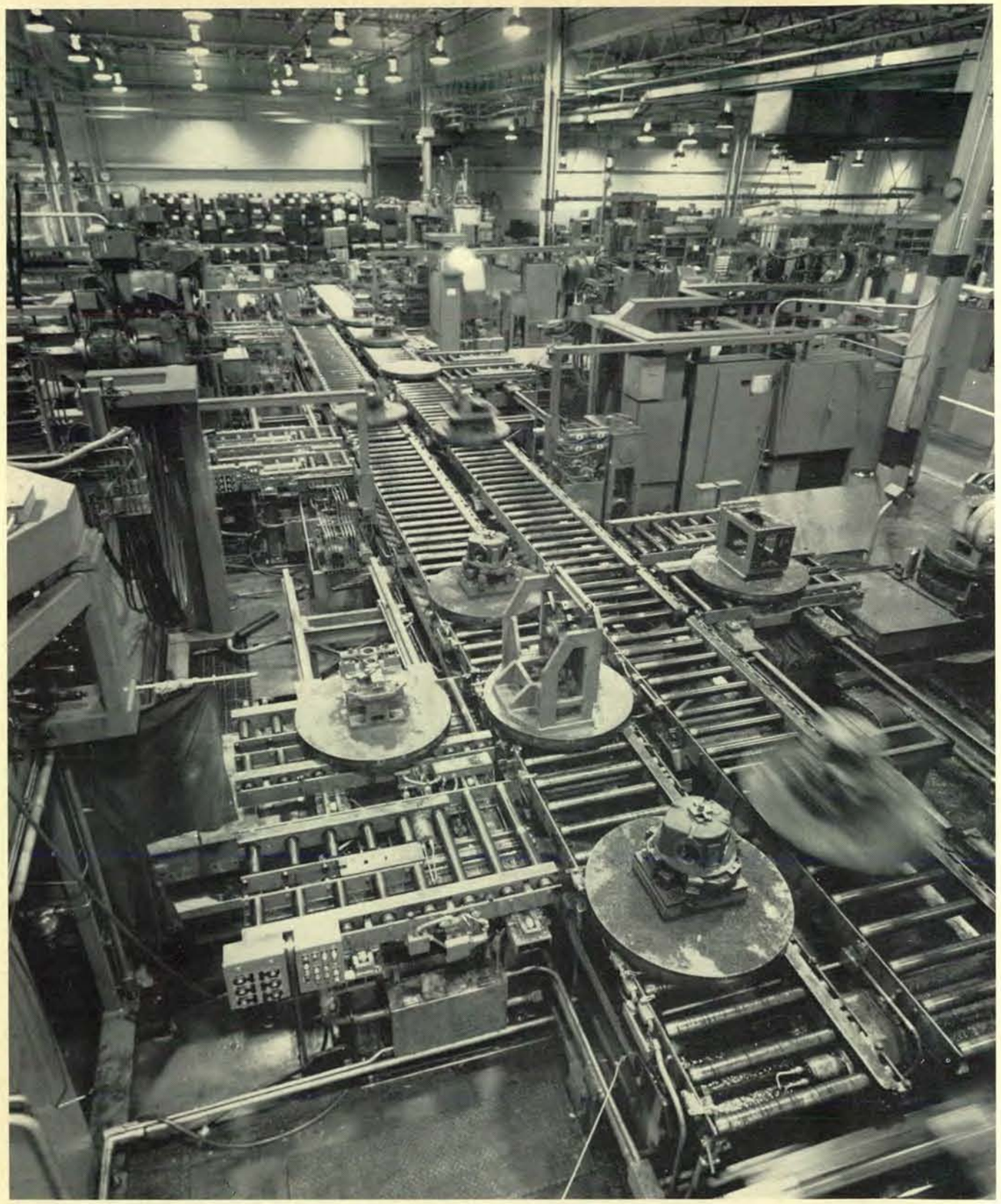


Il Centro di lavorazione pesante nello stabilimento della Ingersoll-Rand Company a Roanoke in Virginia è stato probabilmente il primo esempio importante della fabbricazione di pezzi governata da calcolatore (CMPM) negli Stati Uniti. Il centro, progettato e costruito dalla Sundstrand Corporation, è entrato in funzione nel 1972. Intorno al sistema di trasferimento dei pezzi da una macchina utensile a quella successiva (*si veda il diagramma del ciclo di lavorazione nella pagina a fronte*) sono installate sei macchine utensili a controllo numerico: due fresatrici a cinque assi, due fresatrici a due assi e due trapanatrici a quattro assi. Le sei macchine e il sistema di trasferimento sono controllati da un calcolatore IBM 360/30. A causa dei tre secondi di esposizione necessari per riprendere questa fotografia, l'immagine del pezzo che ha appena iniziato il trasferimento risulta sfocata. Il centro fabbrica una grande quantità di pezzi impiegati nei prodotti della Ingersoll-Rand, principalmente per paranchi e argani. Sui supporti circolari sono fissati i pezzi le cui dimensioni massime non oltrepassino i 590 cm^3. Le lavorazioni comprendono operazioni di fresatura, tornitura, alesatura, maschiatura e trapanatura. Il calcolatore può scegliere fra circa 200 utensili disposti in caroselli e contemporaneamente disponibili per il cambio automatico. Il centro è controllato da 3 operatori e da un supervisore. Per ottenere la stessa produzione un'officina convenzionale avrebbe bisogno di 30 macchine e 30 operai.

LE SCIENZE
edizione italiana di SCIENTIFIC AMERICAN

numero 82
giugno 1975
anno VIII
volume XIV

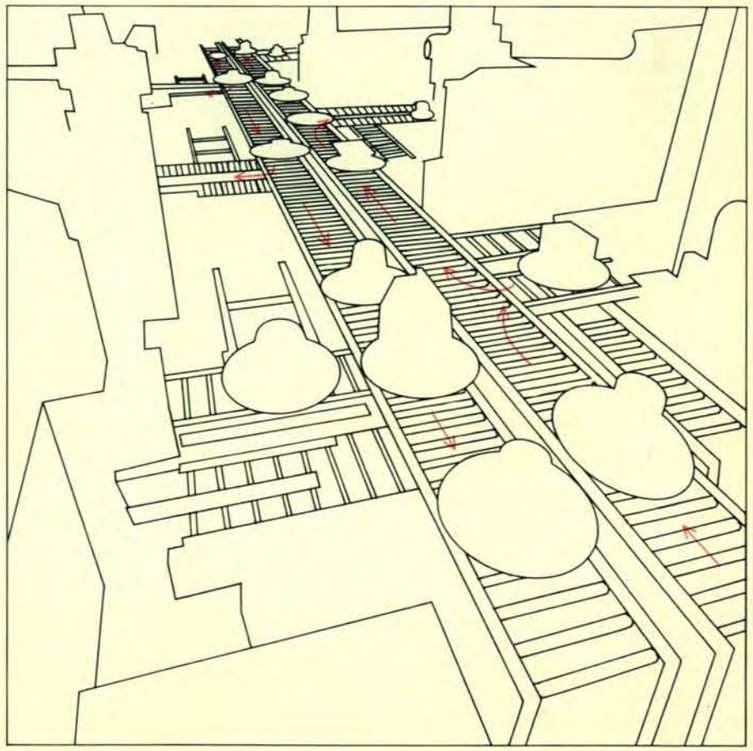
Fabbricazione di pezzi con il calcolatore

Benché non sia stata ancora realizzata una fabbrica completamente automatizzata, sono stati fatti notevoli progressi nella messa a punto di sistemi automatici per la fabbricazione «a lotti» di pezzi complessi

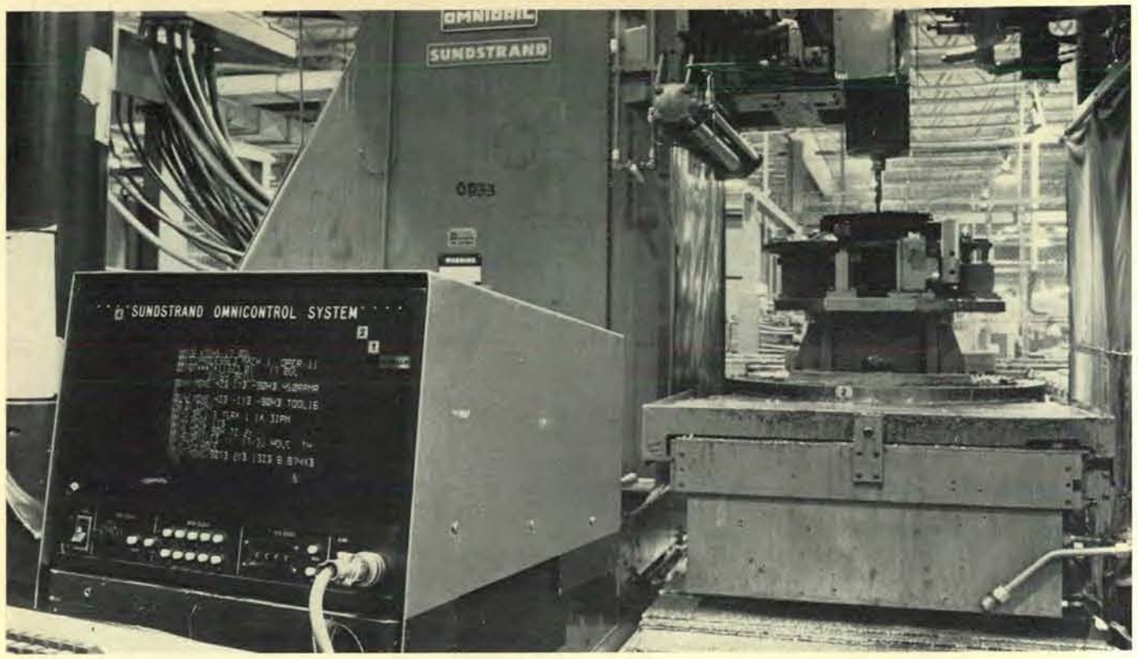
di Nathan H. Cook

L'evoluzione in atto nella tecnologia manifatturiera sta aprendo la porta a un nuovo tipo di fabbrica, nella quale l'ambiente di lavoro sarà molto migliore, il fabbisogno di mano d'opera enormemente ridotto, i materiali grezzi utilizzati in modo più completo e i prodotti fabbricati a un costo notevolmente inferiore. Gli sviluppi a cui si devono questi miglioramenti implicano una applicazione più estesa dei calcolatori per il controllo e l'azionamento dei sistemi e delle macchine operatrici. Poiché gli sviluppi in questo campo sono recenti alcuni termini del problema non sono ancora conosciuti e mi propongo di illustrarli in questo articolo.

Al livello massimo e più avanzato vi sono (o vi saranno) sistemi di fabbricazione integrati con calcolatori (CIMS, dall'inglese *Computer-Integrated Manufacturing Systems*). Tali sistemi richiedono la partecipazione coordinata dei calcolatori a tutte le fasi di lavoro nell'azienda manifatturiera: il progetto del prodotto, la pianificazione della sua fabbricazione, la produzione automatica delle parti, il montaggio automatico, il collaudo automatico e naturalmente il flusso dei materiali e delle parti attraverso l'impianto controllato dai calcolatori. Per quanto si sappia non esistono ancora sistemi di questo tipo completamente realizzati; tuttavia stanno comparendo vari sot-



Nel centro di lavorazione pesante della Ingersoll-Rand i pezzi da lavorare seguono il ciclo qui presentato. Il calcolatore può governare contemporaneamente la fabbricazione di 16 pezzi di tipo diverso. Secondo la complessità del pezzo, esso può essere indirizzato verso due o tre delle sei macchine saltando le altre. Un pezzo normale, per presentare agli utensili tutte le superfici, non deve essere montato più di due volte sul supporto, mentre in una officina convenzionale può essere necessario montarlo da cinque a quindici volte.



Nel Centro di lavorazione pesante della Ingersoll-Rand lo schermo di presentazione a raggi catodici disposto vicino a ognuna delle macchine utensili indica il pezzo in lavorazione e quale fase è stata eseguita dalla macchina nella sequenza delle operazioni stabilite dal calcolatore. L'indicatore che appare nella fotografia è associato in questo caso a una trapanatrice a quattro assi.

Il carosello per il cambio automatico degli utensili che equipaggia ogni macchina Sundstrand nell'impianto Ingersoll-Rand può contenere fino a 60 utensili diversi. Gli utensili sono scelti in sequenza dal programma del calcolatore e vengono fissati automaticamente alla macchina in pochi secondi. Gli utensili sono perpendicolari al carosello anche se nella foto appaiono inclinati.

tosistemi e componenti di questi sistemi. In questa sede descriverò un sottosistema chiamato da me e dai colleghi del Massachusetts Institute of Technology «fabbricazione di pezzi governata dal calcolatore» (CMPM, dall'inglese *Computer-Managed Parts Manufacture*).

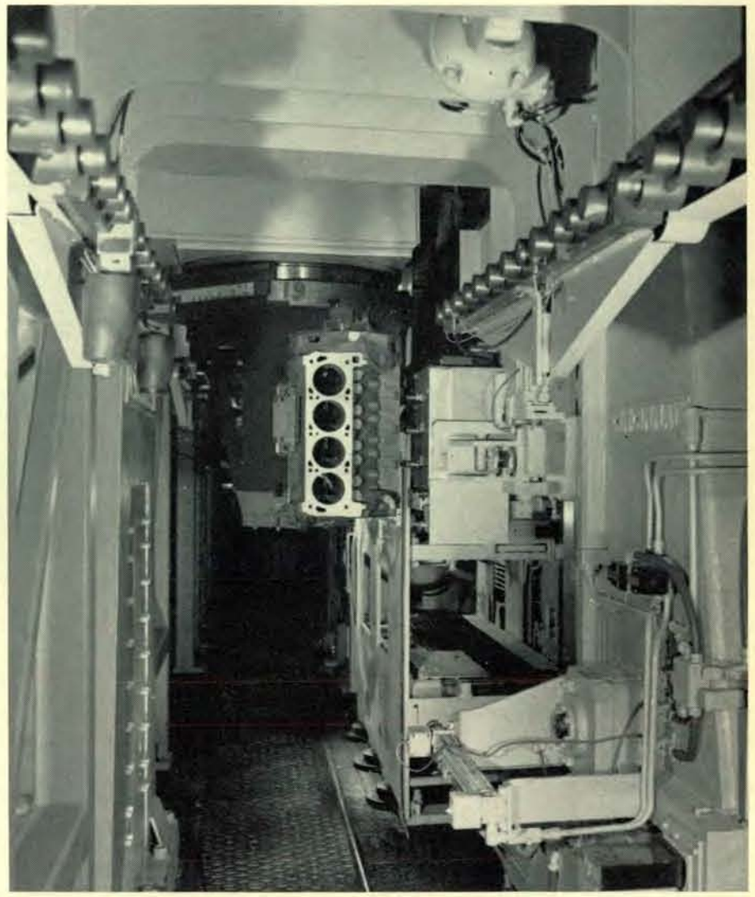
Allo scopo di dimostrare le possibilità della fabbricazione di pezzi governata da calcolatore verranno anzitutto descritti brevemente i metodi attuali di fabbricazione. Per semplicità si prende in considerazione solo la produzione di parti in metallo come quelle impiegate nelle automobili, nelle macchine lavastoviglie, nei frigoriferi, negli apriscatole, nei temperamatite, eccetera. Gli americani sono abituati a pensare alle loro fabbriche come ad aziende ad alta produttività; l'efficienza della produzione in serie fa parte del loro patrimonio nazionale. Fin dalla scuola elementare essi sentono parlare di Henry Ford e sono informati che la loro produttività è la più alta del mondo in termini di produzione di beni per ora di mano d'opera prestata e tendono a credere, con compiacimento, che la loro produttività è forse quasi la massima raggiungibile. Ma ciò non è vero.

Per gli articoli prodotti in serie, ad esempio i motori per automobili, i sistemi americani sono in effetti tanto altamente produttivi che per il futuro oggi prevedibile si può pensare solo a miglioramenti marginali. Però per gli articoli non prodotti in serie, come i motori per aeroplani o le pesanti e costose macchine per centrali elettriche, la produttività è così bassa da essere scoraggiante e quindi la possibilità di realizzare miglioramenti sensazionali è elevata.

La produzione in serie, nell'uso corrente del termine, si riferisce a un particolare gruppo di macchine operatrici progettate di norma per costruire iterativamente lo stesso pezzo nel modo più efficiente possibile. Se si vuole ammortizzare il costo di impianto del gruppo speciale di macchine dedicate alla fabbricazione di questo singolo pezzo è evidente che la produzione annuale deve essere tanto elevata da tenere le macchine quasi ininterrottamente in funzione.

Per molti articoli la domanda non è tale da giustificare la installazione di macchine speciali e quindi devono essere usate macchine per lavorazioni generiche, cioè macchine che possono essere predisposte per realizzare una grande varietà di pezzi. La maggior parte delle persone ha una certa familiarità con le macchine utensili per lavorazioni generiche come i torni, le trapanatrici, le fresatrici e simili. Mediante attrezzi adatti e operai specializzati con una serie di tali macchine si può fabbricare una grande varietà di pezzi, il cui costo però è elevato in confronto a quello dei pezzi fabbricati in grande serie.

Un determinato pezzo può essere realizzato in diversi modi partendo dalla fabbricazione mediante macchine per lavorazioni generiche fino alla fabbricazione con i metodi della produzione di serie. La scelta logica del tipo di macchine da usare dipende principalmente dal numero dei pezzi da eseguire, numero che può variare da una sola unità a milioni di unità all'anno. Normalmente il costo della lavorazione alla macchina è funzione della quantità da produrre. Lavorare una sola unità con macchine utensili generiche può costare fino a 100 volte di più di quanto costa fabbricare lo stesso pezzo con i più efficienti metodi della produzione in serie. Come esempio si consideri un pezzo complesso prodotto in serie, conosciuto da chiunque: il monoblocco di un tipico motore a 8 cilindri a V per automobile. Nelle condizioni della produzione di grande serie, nella quale il monoblocco viene trasferito automaticamente lungo la linea di trasporto da una stazione di lavorazione alla successiva per le varie operazioni (trapanatura, maschiatura, alesatura, fresatura, ecc.), l'intero costo di lavorazione alla macchina (escluso il costo delle materie grezze, ma inclusi i costi della mano d'opera e degli attrezzi, il deprezzamento del macchinario e l'interesse sull'investimento) sarebbe dell'ordine di 25 dollari (16 000 lire). Tuttavia se si dovessero costruire solo alcuni monoblocchi speciali con le macchine



Il sistema di fabbricazione per lavorazioni multiple in fase di sviluppo da parte della Cincinnati Milacron Inc. impiega un sistema non convenzionale di trasferimento sospeso per muovere i pezzi in lavorazione. Nella fotografia un monoblocco viene convogliato verso un «centro di fabbricazione» a cinque assi a controllo numerico che cambia automaticamente le teste portautensili, alcune delle quali portano punte da trapano, maschi per filettare o utensili di vario genere per l'esecuzione di operazioni simultanee.

utensili generiche e operai specializzati, il costo della lavorazione potrebbe facilmente aumentare da 25 a 2500 dollari o più (da 16 000 a 1 600 000 lire). Per produzioni in quantità intermedie in cui i pezzi vengono fatti « per lotti » oppure per partite, è tuttora necessario impiegare le macchine utensili generiche (che possono essere usate anche per fare altri pezzi); si può giustificare, però, un certo grado di automazione con una corrispondente riduzione del costo per pezzo di lavorazione alla macchina.

Si può ora esaminare il problema più da vicino e domandarsi dove viene speso il denaro per la fabbricazione dei pezzi. Sebbene l'industria produca in serie una grande quantità e varietà di pezzi, il loro costo unitario è basso. D'altra parte i pezzi « costruiti in un solo esemplare », sebbene molto costosi, sono anche in numero piuttosto limitato. Fra questi due estremi ci sono pezzi di molti tipi che vengono prodotti in lotti; sia i costi sia la quantità di questi ultimi non sono molto elevati. È stato valutato che negli Stati Uniti fra il 50 e il 75 per cento della spesa complessiva per la fabbricazione di pezzi è relativa ai metodi di fabbricazione a lotti con ordine di grandezza del singolo lotto pari a 50 unità o anche meno. Le parti costruite in questo modo possono costare da 10 a 30 volte di più che se fossero prodotte in serie. Quindi ciò ha una notevole influenza economica: se una parte dell'efficienza della produzione automatizzata di serie potesse essere utilizzata per la fabbricazione a lotti, si avrebbero grandi possibilità di risparmio.

Non deve meravigliare che gli operai addetti alla fabbricazione per lotti delle parti metalliche costituiscano circa il 40 per cento di tutta la forza di lavoro manufatturiera. Diventa sempre più difficile trovare apprendisti per le industrie metalmeccaniche poiché sembra che i giovani preferiscano altri tipi di occupazione, particolarmente nel settore dei servizi. Tale fenomeno si verifica su scala mondiale; molte nazioni europee sono già costrette a importare per l'industria un gran numero di operai semispecializzati e non specializzati. Quindi questa è l'area alla quale ci dedicheremo: il settore dell'industria manifatturiera per la fabbricazione in lotti di parti metalliche.

Di quanto si riducono esattamente i costi di lavoro e di produzione con l'introduzione del controllo e della gestione mediante calcolatore? La risposta può essere più precisa prendendo in considerazione come si sono sviluppate, nell'arco degli ultimi dieci anni, le macchine utensili per lavorazioni generiche mediante l'introduzione della tecnologia dei calcolatori.

Nel funzionamento di una macchina per lavorazioni generiche vi è un certo numero di funzioni che devono essere eseguite o manualmente o automaticamente: 1) trasportare il pezzo da lavorare fino alla macchina; 2) montare il pezzo sulla macchina e fissarlo rigidamente e accuratamente; 3) scegliere l'utensile appropriato e inserirlo nella macchina; 4) stabilire e regolare la velocità e altri parametri di funzionamento della macchina; 5) controllare il movimento della macchina mettendo l'utensile in condizioni di compiere la lavorazione voluta; 6) cambiare in sequenza gli utensili, le condizioni e i movimenti fino a completare tutte le operazioni eseguibili sulla macchina; 7) smontare il pezzo dalla macchina. Durante il ciclo di funzionamento della macchina tradizionale per lavorazioni generiche (il tornio, la fresatrice, il trapano, ecc.) tutte le sette funzioni suddette vengono eseguite dall'operatore (*si veda l'illustrazione nella pagina a fronte in alto a sinistra*).

Due decenni fa veniva sviluppata al MIT la prima macchina utensile a controllo numerico. In sostanza la funzione n. 5, il controllo del movimento della macchina, veniva sottratta all'operatore e resa automatica per mezzo dell'informazione contenuta in un nastro perforato (*si veda l'illustrazione nella pagina a fronte in alto a destra*). Attualmente esistono macchine a controllo numerico con vari gradi di complessità a seconda del numero di assi (movimenti) controllati e se viene verificata solo la posizione finale (funzionamento *point-to-point*) o è prescritta l'esecuzione di un profilo continuo. La macchina più semplice di questo tipo è una « trapanatrice *point-to-point* a due assi » nella quale la tavola portapezzo può essere posizionata automaticamente solo nel piano orizzontale (*x-y*). La macchina più complessa è una « fresatrice a cinque assi » nella quale vengono controllati in modo sincrono e continuo tre movimenti lineari (*x-y-z*) e due rotazioni angolari per produrre pezzi scolpiti con profili complessi. Tali macchine sono largamente usate nella costruzione di cellule per aerei.

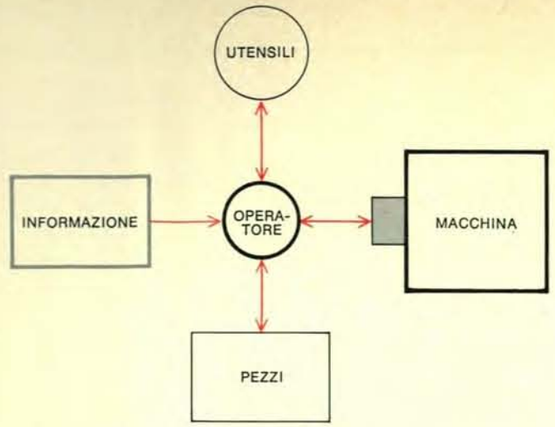
Lo sviluppo delle macchine a controllo numerico ha rappresentato un importante progresso tecnico perché di solito le altre funzioni richiedono molto meno addestramento e abilità da parte dell'operatore. Con l'aumentare del numero di macchine a controllo numerico installate, si è originato un tipo di lavoro assolutamente nuovo e altamente specializzato, quello del programmatore di pezzi la cui funzione è di trasferire il disegno tecnico di un pezzo su nastro perforato il quale poi governa la macchina per produrre il pezzo stesso.

Il risultato più importante che si è avuto dopo quello delle macchine a controllo numerico è stato l'introduzione dei sistemi con cambio automatico dell'utensile (ATC, dall'inglese *Automatic Tool-Changing*), che sottrae all'operatore la funzione n. 3 (*si veda l'illustrazione in basso a sinistra nella pagina a fronte*). Questo tipo di sistema fornisce a una macchina utensile automatica, per esempio una fresatrice a tre assi a controllo numerico, la possibilità di immagazzinare, selezionare e cambiare gli utensili da taglio mediante comandi registrati su nastro perforato. Queste macchine, chiamate spesso centri di lavorazione, hanno raggiunto uno sviluppo che consente a una singola macchina di incorporare una batteria di 60 utensili che possono essere cambiati in pochi secondi.

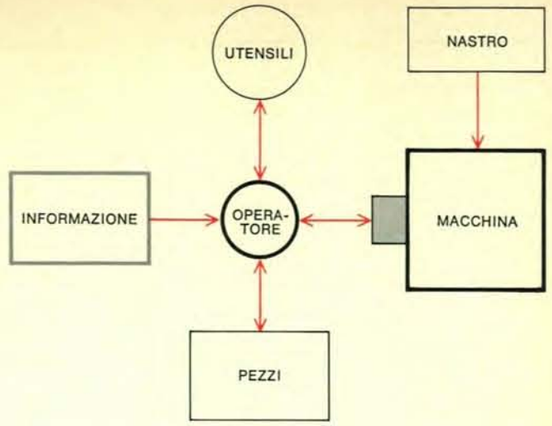
Il passo successivo nello sviluppo delle macchine a controllo numerico è stato la sostituzione dell'informazione contenuta in un nastro perforato con una memoria a dischi magnetici o con organi di memoria per calcolatori dando luogo a un sistema di controllo numerico mediante calcolatore (CNC, dall'inglese *Computer-Numerical-Control*). Questo progresso (*si veda l'illustrazione nella pagina a fronte in basso a destra*) ha reso molto più facile redigere e modificare i programmi per la costruzione delle parti; ha reso inoltre possibile al calcolatore di adempiere a molte funzioni ausiliarie come la preselezione dell'utensile successivo richiesto e averlo già pronto al momento in cui occorre. Quando una batteria di macchine viene posta sotto il controllo di un solo calcolatore ne risulta un sistema conosciuto come controllo numerico diretto (DNC, dall'inglese *Direct Numerical Control*). Ovviamente questo sistema aumenta l'utilizzazione dei calcolatori.

Si possono ora esaminare le caratteristiche e le possibilità di un sistema di fabbricazione di pezzi governato da calcolatori. In un sistema CMPM, un certo numero di macchine sono connesse fra loro non solo tramite un calcolatore di controllo comune, ma anche mediante un sistema di carico/scarico e di trasferimenti dei pezzi (*si veda l'illustrazione a pagina 15*). Oltre alla manutenzione delle macchine e degli utensili l'unica attività dell'operatore è di installare i pezzi da lavorare su supporti speciali (carico) e di rimuovere i pezzi finiti (scarico). Dopo che il pezzo da lavorare è stato installato e ne è stato informato il calcolatore, quest'ultimo ne assume il controllo e avvia il pezzo alle macchine occorrenti, sceglie gli utensili adatti, esegue le opportune operazioni e a lavorazione finita restituisce il pezzo all'operatore per lo scarico. Tutte le sette funzioni richieste sono state automatizzate con il risultato che l'operatore serve un sistema di macchine anziché una sola macchina.

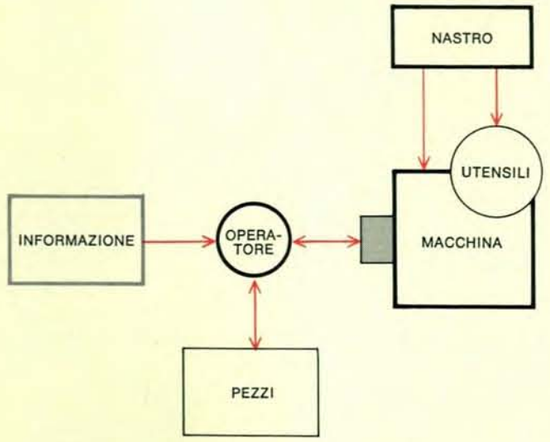
Questo sistema è molto adatto per il settore industriale preso in considerazione: la fabbricazione di pezzi in piccole serie. I pezzi sono prodotti in lotti a causa dei costi della « messa a punto » e di quelli occorrenti per attrezzare la macchina affinché passi dalla fabbricazione del pezzo *A* a quella del pezzo *B*. I dispositivi di fissaggio, che permettono alla parte *A* di essere assicurata perfettamente alla macchina, devono essere sostituiti da altri idonei per la parte *B*; gli utensili per la parte *B* devono essere estratti dal « magazzino degli utensili » e inseriti al posto di quelli usati per la parte *A*; infine la macchina deve essere fatta funzionare e i pezzi prodotti devono essere verificati attentamente per assicurarsi che tutto è in ordine. Ovviamente se viene costruito solo un piccolo lotto di pezzi, i costi unitari di messa a punto sono elevati. Se invece viene costruito un grande lotto i costi di messa a punto sono bassi, ma risultano elevati quelli del successivo magazzinaggio. Ne consegue che esiste una dimensione ottimale del lotto per ogni pezzo particolare. Tuttavia in un sistema di fabbricazione di pezzi governata da calcolatore, il concetto abituale della dimensione economica del lotto non è applicabile perché i costi di



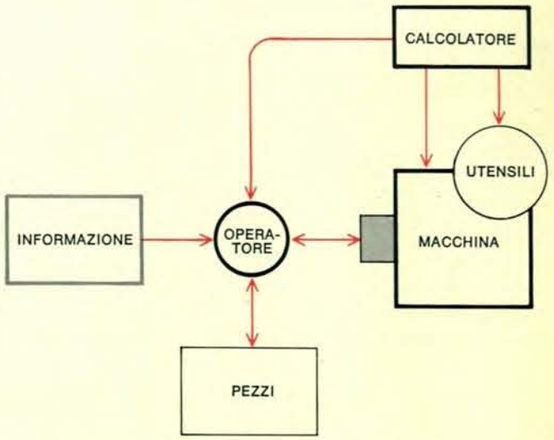
Le macchine per lavorazioni generiche, per esempio un tornio, o una trapanatrice, richiedono le prestazioni di un operatore specializzato. Il suo lavoro consiste nel caricare la macchina, scegliere gli utensili appropriati e provvedere alle necessarie regolazioni della macchina per compiere il lavoro voluto. Di solito le istruzioni sono inserite nel disegno di produzione del pezzo.



La macchina utensile a controllo numerico, sviluppata al Massachusetts Institute of Tecnology, solleva l'operatore dal lavoro di interpretare il disegno di produzione per stabilire le regolazioni della macchina, il cui movimento è controllato dalle istruzioni contenute in un nastro perforato. L'operatore deve tuttavia scegliere ancora gli utensili e caricare e scaricare la macchina.



Il passo successivo per aumentare la produttività delle macchine utensili per lavorazioni generiche è stato il cambiamento automatico degli utensili. In questo sistema il nastro perforato contiene anche le informazioni per scegliere da una batteria di utensili, che varia da 20 a 100, quello appropriato. Il tempo occorrente per il cambio dell'utensile può essere di soli due secondi.



Nel controllo numerico mediante calcolatore l'immagazzinamento dell'informazione è trasferito da un nastro perforato alla più ampia e flessibile memoria di un calcolatore. Questo cambiamento non solo rende più facile compilare e modificare i programmi, ma consente anche di utilizzare il calcolatore per altri compiti come la registrazione della durata d'impiego di ogni utensile.

messa a punto sono estremamente bassi. Può convenire lavorare anche un « lotto » composto di un unico pezzo. Se i lotti sono esigui diventa quasi trascurabile il costo di magazzinaggio di parti finite o in corso di lavorazione.

Non c'è bisogno di mettere in evidenza che i sistemi di fabbricazione di parti governati da calcolatore richiedono una quantità di lavoro diretto considerevolmente minore dei sistemi di fabbricazione più convenzionali. I dati, per quanto ancora limitati, indicano che un sistema CMPM richiede solo dal 10 al 30 per cento del lavoro diretto occorrente in un'officina convenzionale di capacità equivalente. Senza dubbio è anche vero che le macchine per la fabbricazione di parti dirette da calcolatore sono più complesse e costose di quelle convenzionali. Può il lavoro risparmiato compensare il loro maggior costo? Per rispondere a questa domanda si deve ricordare che una macchina per lavorare il metallo è produttiva solo nei periodi in cui il truciolo viene effettivamente asportato, mentre non è produttiva quando è ferma, quando viene attrezzata per produrre un pezzo diverso, quando i pezzi vengono caricati e scaricati o mentre vengono sostituiti gli utensili. Anche se le stime differiscono fra loro, la macchina utensile comune in un'officina convenzionale taglia il metallo solo dal 3 al 10 per cento del tempo, mentre in un sistema CMPM il tempo di taglio del metallo può essere del 50 per cento o anche più. È questa utilizzazione di gran lunga maggiore che compensa il costo più elevato della macchina singola. Si prevede che per una determinata capacità produttiva l'investimento di capitale richiesto per un sistema CMPM altamente sofisticato non sarà maggiore di quello occorrente per un'officina più convenzionale.

In aggiunta ai sicuri risparmi nei costi di lavoro si può contare su notevoli economie per le minori scorte di magazzino. Un numero molto minore di pezzi, sia in lavorazione sia finiti, sarà « in attesa » o della successiva lavorazione o dell'assemblaggio. Come risultato si prevede una riduzione del costo complessivo di un fattore fra cinque e dieci; in altri termini una riduzione di costo dall'80 al 90 per cento.

Sebbene i sistemi CMPM siano solo ora in corso di realizzazione ne sono stati installati alcuni prototipi per la produzione o a scopi dimostrativi. Il sistema forse più ambizioso di tutti (e anche quello sul quale è molto difficile ottenere informazioni) è in funzione a Karl-Marx-Stadt nella Germania Orientale. Il sistema, installato in uno speciale edificio ad aria condizionata grande quanto due campi di calcio, impiega motori a induzione lineare per trasferire i supporti di lavoro, sospesi su cuscini d'aria. Questo sistema può lavorare pezzi con dimensioni massime di 1 × 1 × 1,6 metri.

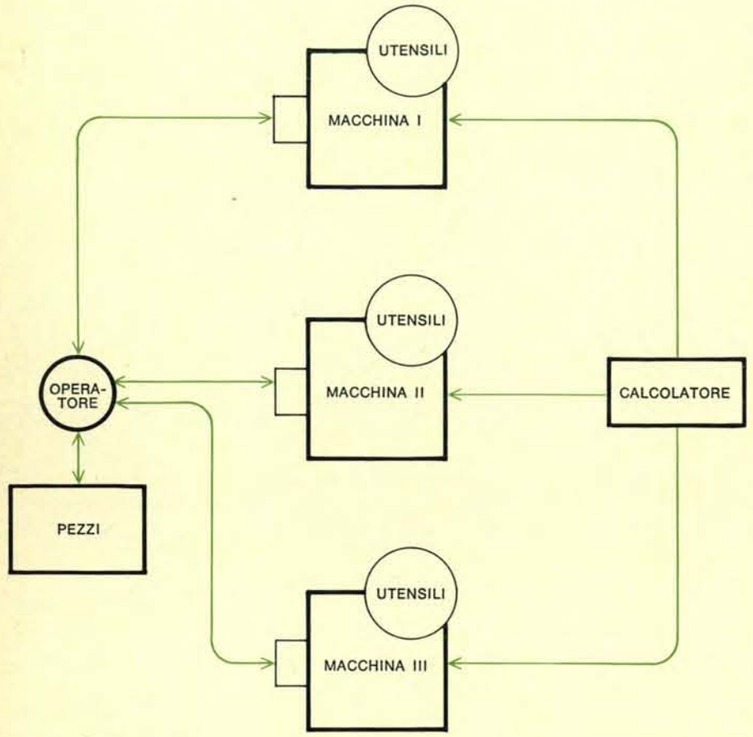
Il sistema CMPM che è stato probabilmente più a lungo in esercizio negli Stati Uniti è il Centro di lavorazione pesante nello stabilimento Ingersoll-Rand a Roanoke in Virginia. Costruito dalla Sundstrand Corporation, l'impianto ha sei macchine disposte intorno a un sistema di trasferimento chiuso (*si vedano le illustrazioni alle pagine 8, 9 e 10*). Il sistema è in grado di fabbricare circa 500 pezzi completamente diversi e può accettare contemporaneamente fino a sedici pezzi di forma diversa sia per essere lavorati alla macchina, sia in attesa di essere lavorati, sia sul sistema di trasferimento. Nell'attuale configurazione il sistema comprende circa 500 utensili, 200 dei quali disposti simultaneamente nei caroselli per il cambio automatico degli utensili. L'intero sistema, che si può considerare come il sostituto di una officina tipica con 30 macchine e 30 operai, è fatto funzionare da una squadra composta da tre persone e da un supervisore.

Alcuni dei più notevoli progressi nella produzione automatica di pezzi in piccole serie sono stati incorporati nel sistema di fabbricazione per lavorazioni multiple attualmente in fase di sviluppo da parte della Cincinnati Milacron Inc. (*si veda l'illustrazione a pagina 11*). Questo sistema ha non solo le caratteristiche generali della fabbricazione di pezzi governata da calcolatore già viste in altri sistemi, ma provvede anche alla lavorazione di pezzi di piccolo volume con un ritmo più alto di quello ottenibile con le macchine convenzionali a controllo numerico. Per esempio, un ingegnoso « centro di fabbricazione » a cinque assi cambia automaticamente gruppi di utensili montati su un'unica testa in modo da effettuare simultaneamente un certo numero di operazioni. Il sistema della Cincinnati Milacron consente, per mezzo di un nuovo metodo per il trasporto sospeso dei pezzi da lavorare, una manovra efficiente del fluido di taglio e dei trucioli, oltre a un facile accesso per le verifiche e la manutenzione.

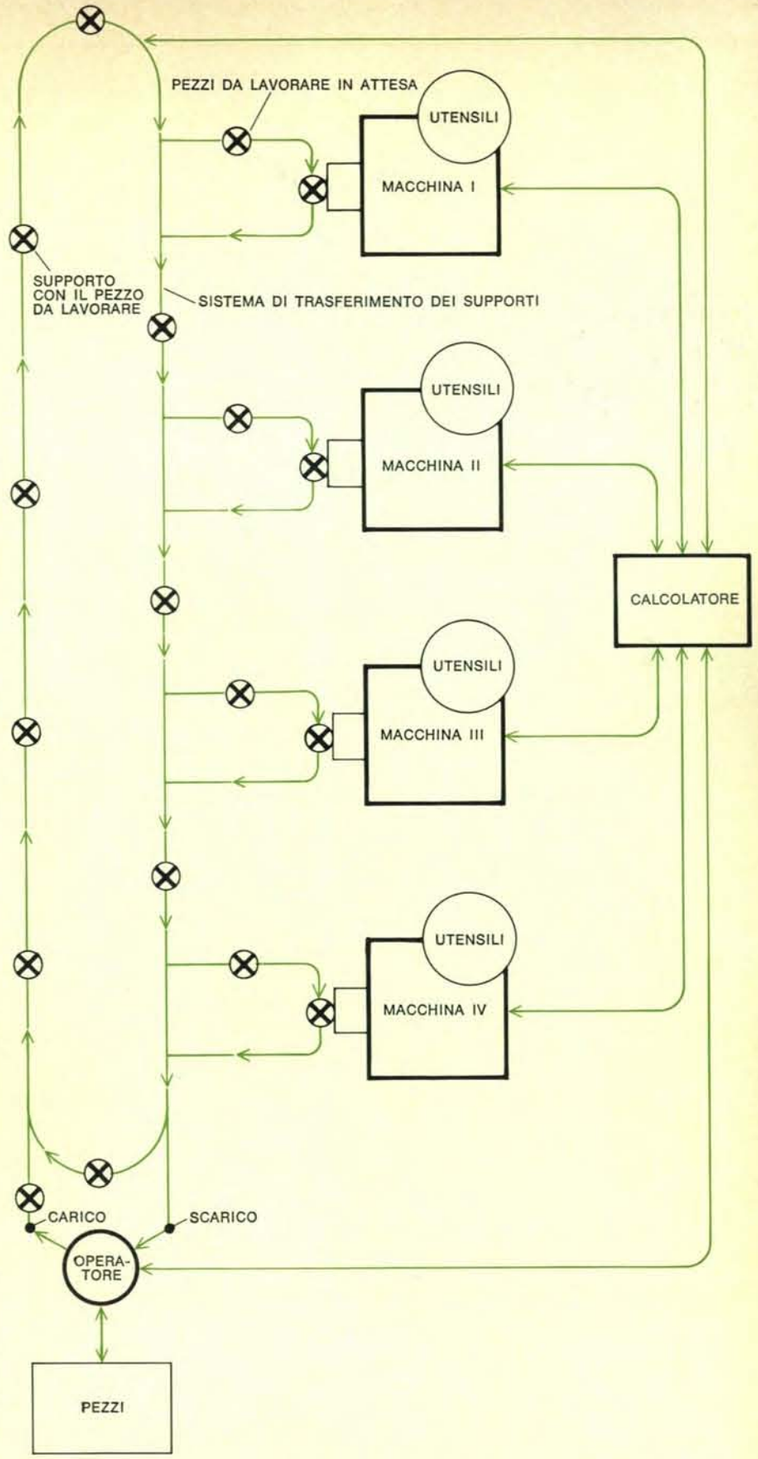
Quanto detto finora si riferisce alla descrizione della fabbricazione di pezzi governata da calcolatore e al suo attuale stadio di sviluppo. Cosa riserba il futuro? A questo punto si devono sostituire le opinioni personali alla descrizione dei fatti. Se per i sistemi CMPM si possono prevedere riduzioni di costi fino all'80 per cento, perché i sistemi in esercizio sono così pochi? I motivi sono di due tipi: tecnologici ed economici.

Il principale motivo economico è che finora non si è completamente sicuri di quanto si possa risparmiare nelle varie branche dell'industria con la fabbricazione di pezzi governata da calcolatore. Come si è visto, si valuta che per il 50-75 per cento tutti i pezzi fabbricati alla macchina costituiscono lotti di 50 unità o meno e quindi adatti a essere prodotti con la fabbricazione governata da calcolatori. Comunque non si conosce quale sia l'incidenza del fattore economico. I sistemi esistenti sono stati giustificati per la fabbricazione di specifiche serie di pezzi e la giustificazione ha incluso una notevole quantità di fiducia. I sistemi CMPM sono dispendiosi: possono costare milioni di dollari. A meno che non si possa provare con la ricerca e le dimostrazioni che i sistemi CMPM sono davvero economici, affidabili e suscettibili di future modifiche nella varietà e quantità di prodotti, poche società saranno disposte ad assumere il rischio finanziario richiesto per scoprirlo in proprio.

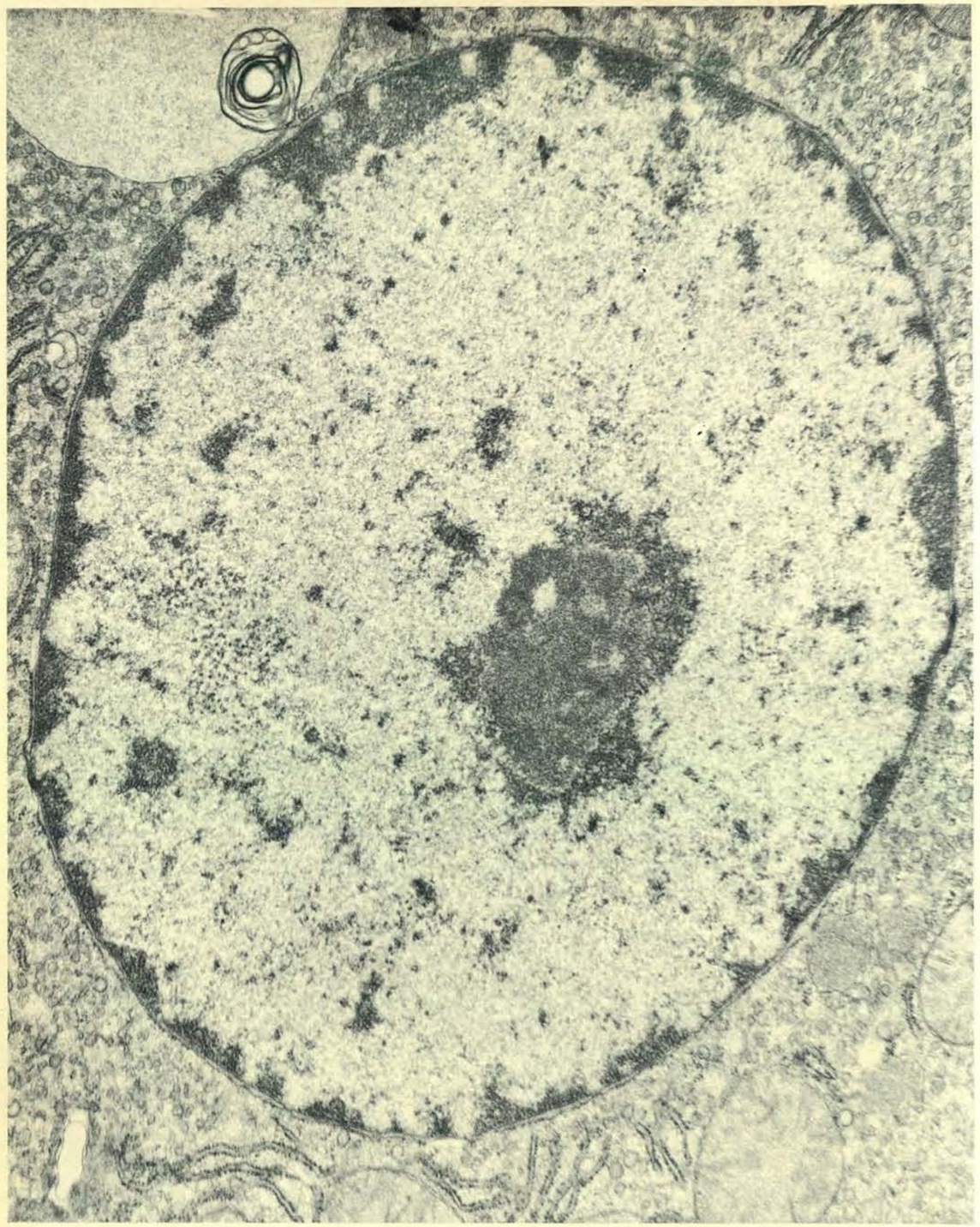
I motivi tecnologici della lenta adozione dei sistemi CMPM sono soprattutto collegati allo sviluppo dei sistemi di calcolatori e dei programmi per il loro funzionamento. Quantunque singole società abbiano creato sistemi funzionanti in modo soddisfacente, per quanto si sappia nessuno di questi ha la completa versatilità operativa necessaria per garantire la massima prestazione da parte di un sistema CMPM. Dal punto di vista del mercato degli Stati Uniti sembra troppo dispendioso per ciascun fabbricante di macchine utensili che pensi di entrare nel mercato CMPM sviluppare completamente un proprio sistema brevettato. È opinione, quindi, dell'autore che se i costi valutati del CMPM sono sufficientemente esatti e se si realizzano le speranze degli Stati Uniti di mantenere la produzione a livelli più alti di quelli delle altre nazioni, si dovrà preparare un programma nazionale per lo sviluppo della fabbricazione di pezzi governata da calcolatori con lo scopo finale di realizzare sistemi di fabbricazione totalmente integrati da calcolatori.



Il controllo numerico diretto pone sotto il controllo di un unico calcolatore una serie di macchine, che possono essere di tipo diverso e programmate per lavori differenti. Il sistema può avere o non avere la possibilità del cambio automatico degli utensili.



La fabbricazione di pezzi governata da calcolatore è la fase più recente dell'applicazione della tecnologia dei calcolatori per ridurre il costo della lavorazione di pezzi in piccole serie. Siccome la memoria del calcolatore contiene tutte le informazioni necessarie per lavorare un pezzo alla macchina utensile, un « lotto » può essere composto anche di una sola unità. Il sistema qui descritto è sostanzialmente lo stesso di quello impiegato nel Centro di lavorazione pesante della Ingersoll-Rand di Roanoke in Virginia.



Microfotografia elettronica di nucleo di cellula epatica di topo, effettuata da David E. Comings e Tadashi Okada del City of Hope Medical Center. Il nucleo, delimitato da una doppia membrana, è riempito di cromatina, il materiale che svolge le funzioni genetiche e di controllo della cellula. Nella cromatina, il materiale genetico (DNA) forma un complesso con le proteine la cui funzione è apparentemente quella di regolare l'attività dei geni. Il materiale scuro vicino alla membrana e in altri punti della cellula è cromatina «addensata», che sta sintetizzando RNA. Il materiale granulare fine che è visibile nella metà sinistra del nucleo è costituito da granuli di RNA riuniti per formare un complesso con le proteine. La grossa struttura scura, visibile vicino al centro, è il nucleolo, una zona di cromatina specializzata. Il preparato risulta ingrandito 22 000 volte nella microfotografia.

Proteine dei cromosomi e regolazione genica

La funzione delle proteine nei nuclei degli organismi superiori sta per essere chiarita. Sembra che gli istoni mantengano i geni disattivati e le proteine non istoniche li attivino selettivamente

di Gary S. Stein, Janet Swinehart Stein e Lewis J. Kleinsmith

Come è ben noto, i geni che trasmettono l'informazione ereditaria di generazione in generazione e dirigono il funzionamento di ogni cellula vivente sono costituiti da DNA. Nelle cellule degli organismi superiori i geni sono disposti sui cromosomi, ma questi sono in realtà nucleoproteine, cioè complessi di acidi nucleici e di proteine. Se il materiale genetico è soltanto DNA, qual è allora la funzione delle proteine cromosomiche? Sembra che esse svolgano un compito importante nel mantenere costante la struttura del materiale genetico e nel regolare l'attività dei geni, cioè nel determinare quali geni in ogni cellula vengono attivati e quando. Il controllo da parte dei geni è un fatto chiave in processi come il differenziamento, lo sviluppo embrionale e l'azione ormonale e in processi abnormi come la proliferazione cancerosa, le malattie metaboliche e quelle malattie congenite che sono correlate all'espressione dell'informazione genetica. In questa sede intendiamo interessarci in primo luogo dell'attività regolatrice delle proteine cromosomiche, un campo d'indagine che ha cominciato a fornire risultati significativi negli ultimissimi anni.

L'attività regolatrice dei geni

L'informazione genetica di una cellula è codificata nelle sequenze nucleotidiche che costituiscono il DNA. Per utilizzarla, la cellula trascrive tali sequenze in filamenti complementari di RNA che sono poi «tradotti» in catene di amminoacidi, che a loro volta formano le proteine. Il controllo della attività genica risiede nella selettività della trascrizione sia nel tempo, sia nello spazio. Nello sviluppo embrionale, per esempio, un singolo uovo fecondato contenente tutti i geni dell'organismo prolifera così da dare un'ampia varietà di cellule differenziate, che sono specializzate per compiere diverse funzioni. Nel corso dello sviluppo centinaia di migliaia di geni vengono attivati e disattivati. Inoltre in ogni tipo di cellula definitiva viene trascritta, o espressa, solo una certa combinazione di geni, in base alla funzione specifica di quel tipo cellulare: nei globuli rossi può trattarsi dell'attività del gene che codifica per l'emoglobina; nelle cellule muscolari di quella del gene per la mioglobina; nei fibroblasti di quella del gene per il collagene, la proteina del tessuto connettivo. E pur tuttavia, per quanto si può dire, il DNA è sempre lo stesso in ognuno di questi tipi di cellule specializzate. La costanza del DNA in tutte le cellule di un determinato organismo è stata stabilita per la prima volta nel 1948 ed è stata confermata ripetutamente. La verifica più elegante è venuta dagli esperimenti di J.B. Gurdon dell'Università di Oxford. Questo ricercatore ha trapiantato nuclei di cellule intestinali differenziate di rana in uova di rana i cui nuclei erano stati rimossi: le uova si sono sviluppate normalmente. Risultati del genere dimostrano che i nuclei delle cellule intestinali e di tutte le altre cellule differenziate della rana contengono l'informazione genetica totale della rana, anche se soltanto una porzione limitata di essa viene espressa.

Anche all'interno di una data cellula differenziata, una certa parte dell'informazione si esprime in un determinato momento e rimane inespressa in altri. L'espressione si può modificare quando le cellule in riposo sono attivate per dividersi oppure quando sono stimolate da ormoni specifici; in ambedue i casi si effettuano nel metabolismo della cellula complessi cambiamenti, in seguito all'innesco del programma genetico insito nella cellula stessa. In generale, in qualsiasi momento, si esprime meno del 10 per cento del totale dell'informazione genetica. Meccanismi regolatori specifici attivano e disattivano per la trascrizione regioni particolari del genoma, cioè una intera serie di geni, a mano a mano che cambiano le necessità della cellula.

Nei batteri, che mancano di un nucleo e in cui il genoma è considerevolmente meno complicato di quello degli organismi superiori, è stato compiuto un progresso sostanziale verso il chiarimento del meccanismo della regolazione. Nel colibacillo *Escherichia coli* si è dimostrato che la trascrizione delle singole sequenze genetiche in RNA è governata dal fatto che regolatori specifici si legano e si distaccano in determinati siti del DNA. Nelle cellule delle piante e degli animali superiori il materiale genetico è più complesso, da un punto di vista sia strutturale, sia funzionale. Il DNA è contenuto all'interno di un nucleo delimitato da membrana e, in questa sede, è unito a proteine e a un piccolo quantitativo di RNA per formare quella sostanza complessa che si chiama cromatina o, nella sua forma più organizzata, cromosoma. Finora, in queste cellule dotate di nucleo, o eucariote, non sono state ancora identificate proteine regolatrici specifiche, ma l'evidenza depone in favore del ruolo di elementi regolatori delle proteine cromosomiche.

Gli istoni

Da un punto di vista storico, le proteine cromosomiche sono state suddivise in due gruppi: gli istoni e le proteine non istoniche. Gli istoni sono definiti nel modo migliore dalle loro proprietà chimiche: si tratta di proteine con una carica positiva, riccamente dotate di amminoacidi basici arginina e

lisina e completamente prive dell'amminoacido triptofano. A causa della loro natura basica, possono essere facilmente estratte con acido cloridrico o solforico diluito. Si possono frazionare in cinque classi principali, la maggior parte delle quali è presente in tutte le cellule eucariote. Gli istoni furono scoperti alla fine del XIX secolo, ma fu solo nel 1943 che Edgar e Ellen Stedman, che allora lavoravano all'Università di Edinburgo, attirarono l'attenzione su di essi come possibili regolatori. I dati degli Stedman fecero intravvedere la possibilità che i tessuti in fase di attivo accrescimento contenessero meno istoni dei tessuti che si trovavano in altre fasi; i due ricercatori ritennero quindi che gli istoni funzionassero da repressori, o inibitori, dell'attività biologica. Essi ebbero ragione di postulare una funzione inibitrice per gli istoni, ma la loro conclusione era basata su misure del contenuto istonico che in seguito sono risultate imprecise; oggi sappiamo che il contenuto in istoni delle cellule attive e inattive è identico.

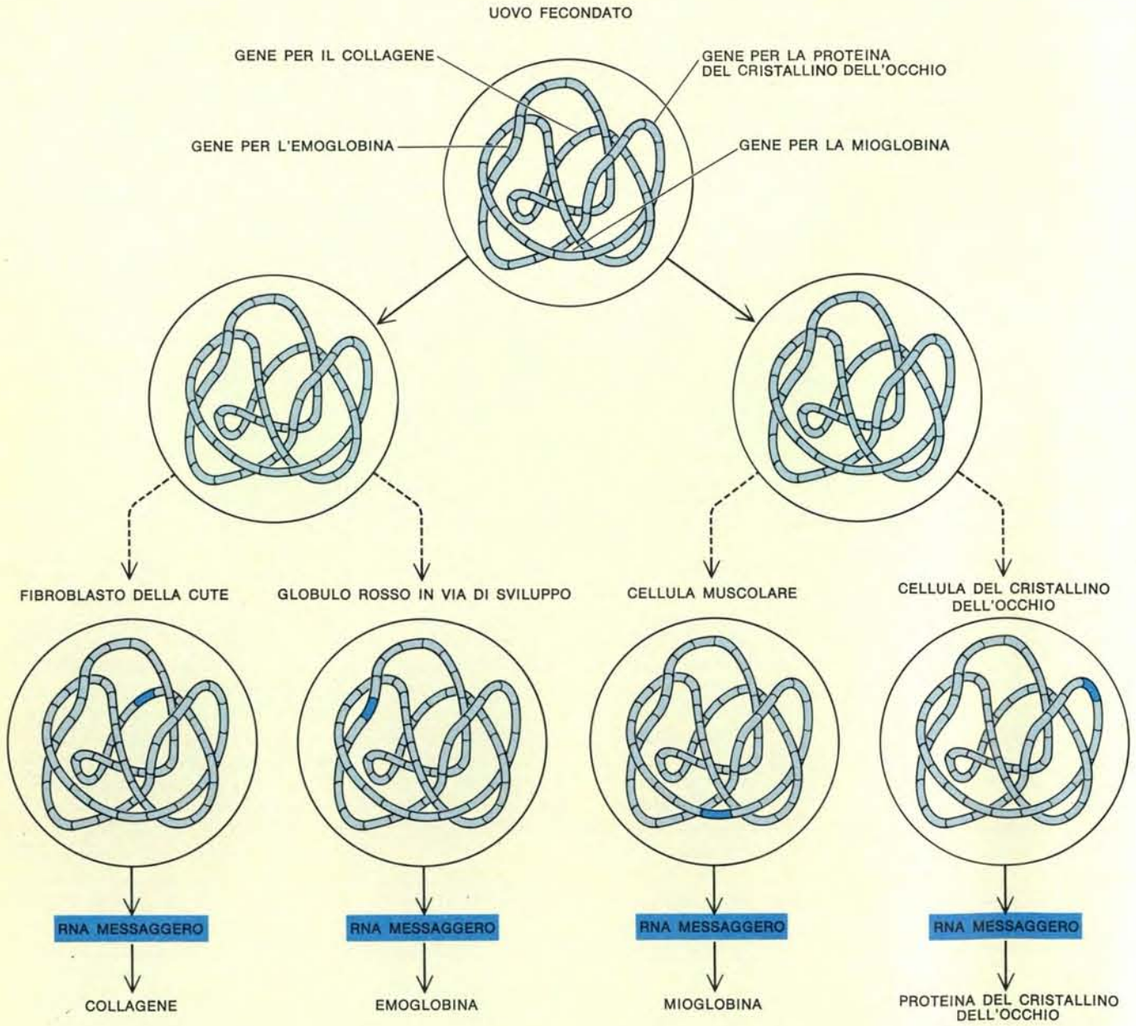
I primi studi decisivi di biochimica sugli effetti degli istoni a carico della funzione del DNA sono stati intrapresi agli inizi degli anni sessanta, dopo che furono sviluppati sistemi acellulari per la sintesi dell'RNA. In tali sistemi viene fornito DNA o cromatina che fungono da stampo per la trascrizione in provetta dell'RNA. Altri componenti della miscela di reazione sono i precursori delle subunità dell'RNA e la RNA polimerasi, l'enzima che catalizza la polimerizzazione di tali subunità nei filamenti dell'RNA. Nel 1962, Ru-chih Huang e James Bonner del California Institute of Technology hanno dimostrato che, aggiungendo istoni a un simile sistema acellulare, si inibisce la sintesi dell'RNA (la massima inibizione della sintesi dell'RNA è stata osservata con un rapporto istone/DNA di 1:1:, che è pressappoco il rapporto in cui i due composti chimici si trovano normalmente nel nucleo). Il processo di inibizione diviene reversibile per aggiunta di un altro DNA, il che dimostra che gli istoni bloccano la trascrizione legandosi al DNA piuttosto che inibendo la RNA polimerasi. All'incirca nella stessa epoca, Vincent G. Allfrey e Alfred E. Mirsky della Rockefeller University hanno dimostrato che la rimozione selettiva degli istoni da nuclei isolati aumenta la velocità della sintesi di RNA. Ambedue i tipi di esperimenti hanno portato alla conclusione che gli istoni normalmente funzionano inibendo la capacità dei geni di essere trascritti nell'RNA.

È subito parso evidente che gli istoni sono implicati anche nella determinazione delle proprietà strutturali della cromatina. In particolare la loro aggiunta aumenta la stabilità della doppia elica del DNA. Questo fatto fornisce una chiave su una delle funzioni biologiche degli istoni, in quanto il serrarsi della doppia elica potrebbe interessare la disponibilità, o esposizione, delle sequenze nucleotiche che vengono trascritte in RNA.

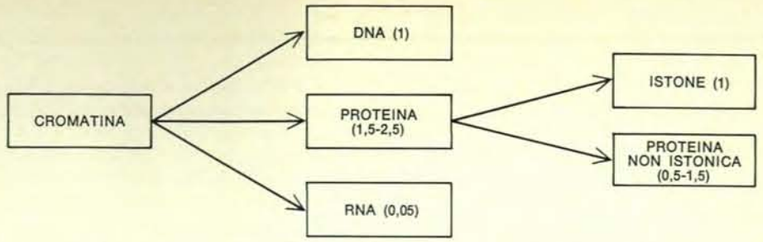
Un altro aspetto del metabolismo degli istoni che fa pensare alla loro funzione biologica è il rapporto tra sintesi degli istoni e replicazione del DNA. Prima che una cellula si divida in due cellule figlie, tutto il suo DNA deve essere duplicato in modo che ogni cellula figlia possa ricevere una copia completa dell'informazione genetica. La sintesi del DNA è limitata a un periodo ben definito del ciclo vitale della cellula (la fase *S*) ed è strettamente accoppiata alla sintesi degli istoni. Thaddeus Borun, Elliot Robbins e Matthew D. Scharff dell'Albert Einstein College of Medicine e Gerald C. Mueller della Università del Wisconsin hanno dimostrato che istone e DNA sono sintetizzati simultaneamente e che la inibizione della replicazione del DNA è accompagnata da un immediato arresto della sintesi dell'istone. Lo stretto appaiamento della sintesi del DNA con quella dell'istone trova appoggio anche in esperimenti che indicano come gli RNA messaggeri che codificano per gli istoni siano associati agli organelli cellulari deputati alla sintesi proteica e tradotti in proteine solo nel momento in cui viene sintetizzato il DNA. (Ciò significa che, se soltanto limitate regioni del genoma devono essere trascritte in una data cellula e in un dato momento, gli istoni devono essere immediatamente disponibili per poter reprimere i segmenti del DNA appena replicati, che devono rimanere inespressi.)

A parte la sintesi di nuove molecole di istone, l'interazione degli istoni con il DNA può essere alterata dalla modificazione di istoni già esistenti. Vi sono reazioni enzimatiche che modificano la struttura istonica mediante la aggiunta di gruppi acetato, metile e fosfato. L'esatto significato di queste reazioni non è ancora noto, ma è stato dimostrato da Allfrey e collaboratori che l'acetilazione dell'istone è correlata al livello generale della sintesi di RNA; la fosforilazione di certi istoni, come ha osservato G. Roger Chalkley dell'Università dello Iowa, è correlata alla velocità di divisione della cellula. Tali reazioni possono essere importanti nel controllo dell'attività globale e della struttura del genoma.

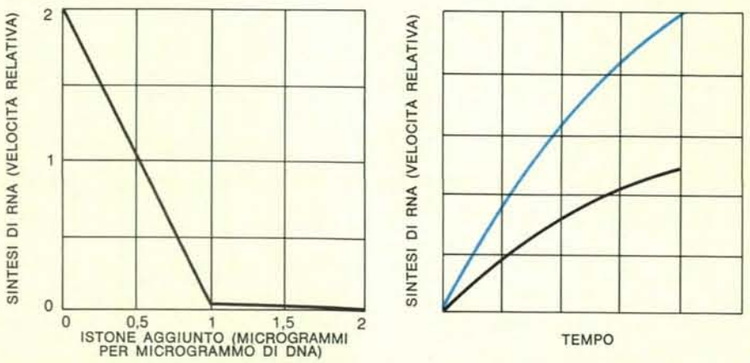
Tutto sommato, è chiaro che gli istoni sono implicati nel mantenimento della struttura cromatinica come pure nella repressione della sintesi di RNA DNA-dipendente. Cionondimeno, parecchie serie di prove suggeriscono che gli istoni hanno un'uniformità e una conseguente mancanza di specificità, che precludono la loro capacità di riconoscere e di influire su particolari geni. In una data cellula analoghe quantità di istoni sono inoltre presenti nelle regioni attive e inattive del genoma, cioè nella cromatina «distesa», che sta sintetizzando attivamente RNA, e nella cromatina «condensata», che non sta sintetizzando RNA. Nei nuclei di certe cellule di larve d'insetto, le regioni del genoma che stanno effettuando la sintesi dell'RNA sono distese a costituire degli ammassi (*puff*) relativamente voluminosi. Hewson Swift dell'Università di Chicago ha misurato il rapporto istone/DNA nelle regioni inattive e in quelle con *puff* e ha trovato che è sempre lo stesso. Forse la prova più sorprendente della mancanza di specificità degli istoni è data dal confronto effettuato tra le sequenze amminoacidiche di una stessa frazione istonica in due tipi molto diversi di cellule, quelle delle pianticelle di pisello e quelle del timo di vitello: è risultato che esse differiscono soltanto per due amminoacidi. Sembra chiaro dunque che gli istoni sono molecole regolatrici implicate nel control-



Viene qui esemplificata la regolazione genica nel corso dello sviluppo. L'uovo fecondato (*in alto*) contiene tutta l'informazione genetica dell'organismo, codificata nel DNA; si tratta dei geni per le quattro importanti proteine qui indicate e per migliaia di altre proteine. L'uovo si divide e le cellule proliferano e si differenziano (*frecce tratteggiate*) in cellule specializzate appartenenti a vari tessuti. Ogni cellula differenziata contiene tutti i geni dell'organismo, ma questi sono regolati in modo che in ogni cellula solo quelli adatti per le funzioni specifiche sono attivati, trascritti in RNA messaggeri e quindi tradotti in proteine.



La cromatina, il complesso materiale presente nel nucleo delle cellule degli organismi superiori, è costituita da DNA, da proteine e da una piccola quantità di RNA. Vi sono due categorie di proteine: gli istoni e le proteine non istoniche. Le cifre indicano le quantità relative di ogni costituente della cromatina. Questa contiene all'incirca la stessa quantità di istoni e di DNA mentre il rapporto con le proteine non istoniche è diverso.



L'effetto inibitorio degli istoni sull'attività dei geni è stato dimostrato per la prima volta negli anni sessanta. Ru-chih Huang e James Bonner del California Institute of Technology hanno trovato (*a sinistra*) che in sistemi acellulari gli istoni inibiscono la sintesi dell'RNA. Vincent G. Allfrey e Alfred E. Mirsky della Rockefeller University hanno dimostrato (*a destra*) che la velocità con cui procede la sintesi dell'RNA è maggiore nei nuclei privati degli istoni (*in colore*) rispetto ai nuclei di controllo (*in nero*).

lo della trascrizione genica. Tuttavia devono esserlo in un senso non specifico.

Proteine non istoniche

Nella ricerca di molecole con capacità di regolazione di geni specifici, l'attenzione è stata puntata di recente sulle proteine non istoniche dei cromosomi, che possono essere definite semplicemente come proteine associate con la cromatina, diverse dagli istoni. Come questi ultimi, esse vengono sintetizzate al di fuori del nucleo, nel citoplasma. Mentre poi gli istoni, subito dopo essere stati sintetizzati, vengono trasportati nel nucleo e associati con il DNA, studi effettuati nei nostri laboratori presso il collegio di medicina dell'Università della Florida e presso l'Università del Michigan hanno dimostrato che, per le proteine non istoniche, il quadro è un poco più complesso. Una parte di queste proteine si associa con il DNA subito dopo la sintesi, ma trascorrono vari periodi di tempo prima che altre proteine non istoniche appaiano tra i componenti della cromatina. Come gruppo, le proteine non istoniche mostrano una velocità di ricambio superiore a quella degli istoni, con una notevole variazione della durata delle singole frazioni non istoniche: alcune hanno una vita di minuti e altre sono stabili perlomeno come il DNA cellulare e gli istoni. Non è chiaro se, abbandonando la loro associazione con il DNA, esse sono degradate oppure entrano a far parte di un fondo comune di proteine nel nucleo o nel citoplasma. Ciò che è chiaro è che, se da una parte sembra che gli istoni e il DNA costituiscano componenti permanenti del genoma, le proteine cromosomiche non istoniche sono, perlomeno in parte, in stato di flusso dinamico.

In contrasto con gli istoni, le proteine non istoniche mostrano un'enorme eterogeneità sia dal punto di vista strutturale che funzionale. Il loro peso molecolare varia tra 10 000 e 150 000 dalton. È difficile determinare accuratamente quante specie di proteine non istoniche sono associate con il genoma, ma parecchie serie di prove fanno pensare che questo numero sia estremamente ampio. Un'indicazione della loro diversità funzionale è la varietà dei sistemi enzimatici complessi che si trovano tra loro. Si hanno enzimi come le polimerasi implicate nella sintesi e nella reintegrazione del DNA e dell'RNA, enzimi implicati nella sintesi e nella degradazione di proteine ed enzimi che modificano gli acidi nucleici e le proteine, aggiungendovi gruppi acetato, metile e fosfato. Oltre a questi componenti enzimatici si ritiene anche che la frazione non istonica contenga proteine che hanno, nella cromatina, funzioni strutturali e, alla stessa stregua, funzioni regolatrici. Nella restante parte di questo articolo concentreremo la nostra attenzione sulle principali serie di prove a sostegno dell'idea secondo cui almeno alcune delle proteine non istoniche sono implicate nella regolazione dell'attività di geni specifici.

Esistono differenze significative tra i tipi di proteine non istoniche che si trovano nelle cellule di differenti specie. Dato che l'informazione genetica codificata nel DNA è unica per ogni specie, l'osservata specie-specificità delle proteine non istoniche è coerente con l'idea che esse abbiano una funzione regolatrice. All'interno dello stesso organismo esistono poi, in tessuti diversi, differenti proteine non istoniche. Anche se tutte le cellule dell'organismo contengono lo stesso DNA, la variazione che si riscontra nell'insieme di proteine non istoniche presenti nei vari tipi cellulari concorda con un'utilizzazione selettiva dell'informazione genetica da parte di quei tipi cellulari.

Se le proteine non istoniche sono impegnate in una regolazione genica specifica, allora la gamma di queste proteine in una cellula dovrebbe modificarsi quando si modifica l'attività dei geni. Quest'ipotesi è stata messa alla prova in un vasto numero di sistemi sperimentali, in cui si verificano notoriamente variazioni dell'espressione genica. In quasi tutti i casi sono stati osservati i cambiamenti previsti nelle proteine non istoniche.

Variazioni nella sintesi

Variazioni nell'espressione genica si verificano, per esempio, nel corso del ciclo cellulare. Le differenze più pronunciate nella trascrizione si osservano tra la fase *S* (il periodo della sintesi di DNA) e la mitosi (il processo di divisione cellulare). Nei nostri e in altri laboratori sono state esaminate durante ogni fase del ciclo cellulare, sotto l'aspetto della composizione e del metabolismo, le proteine cromosomiche non istoniche associate con il genoma e sono state osservate differenze ugualmente pronunciate. Per fare un altro esempio, gli ormoni steroidei, come gli estrogeni e l'idrocortisone, esercitano notoriamente i loro effetti su cellule bersaglio legandosi al loro nucleo e, in ultima analisi, alterando il quadro della trascrizione genica. È stato dimostrato che parecchi ormoni steroidei provocano dei cambiamenti nella composizione e nel metabolismo delle proteine cromosomiche non istoniche.

Le colture in vitro di cellule di neuroblastoma (cancro delle cellule nervose) di topo costituiscono un interessante sistema modello per lo studio del differenziamento. Esse sono facilmente indotte a differenziarsi in cellule con proprietà strutturali, biochimiche ed elettriche tipiche della cellule nervose. Uno dei vantaggi davvero singolari del sistema è che il differenziamento indotto è completamente reversibile. Michelle Zornetzer, nel nostro laboratorio in Florida, ha osservato che, nello stato differenziato, c'è una diminuzione selettiva nella sintesi delle proteine non istoniche di elevato peso molecolare e che questo fenomeno diventa reversibile quando lo diventa anche lo stato di differenziamento (*si veda la figura a pagina 22*). La Zornetzer ha anche dimostrato che il differenziamento delle cellule suddette è associato a differenze nella sintesi dell'RNA. Vi è cioè una forte correlazione tra i tipi di proteine non istoniche associate con il genoma e le alterazioni nell'espressione genica che si accompagnano al differenziamento.

È possibile separare la cromatina in due frazioni che differiscono una dall'altra sia per struttura, sia per attività genetica: la frazione di cromatina condensata, che è strutturalmente compatta e relativamente inattiva nella sintesi dell'RNA, e la frazione di cromatina distesa, che sintetizza attivamente l'RNA. Come abbiamo ricordato, il rapporto istone/DNA è identico in questi due tipi di cromatina. Tuttavia la quantità relativa di proteina non istonica è considerevolmente più elevata nella forma attiva, distesa, della cromatina. Qui di nuovo la presenza delle proteine non istoniche è correlata con uno stato strutturale della cromatina, caratterizzato da una sintesi attiva di RNA.

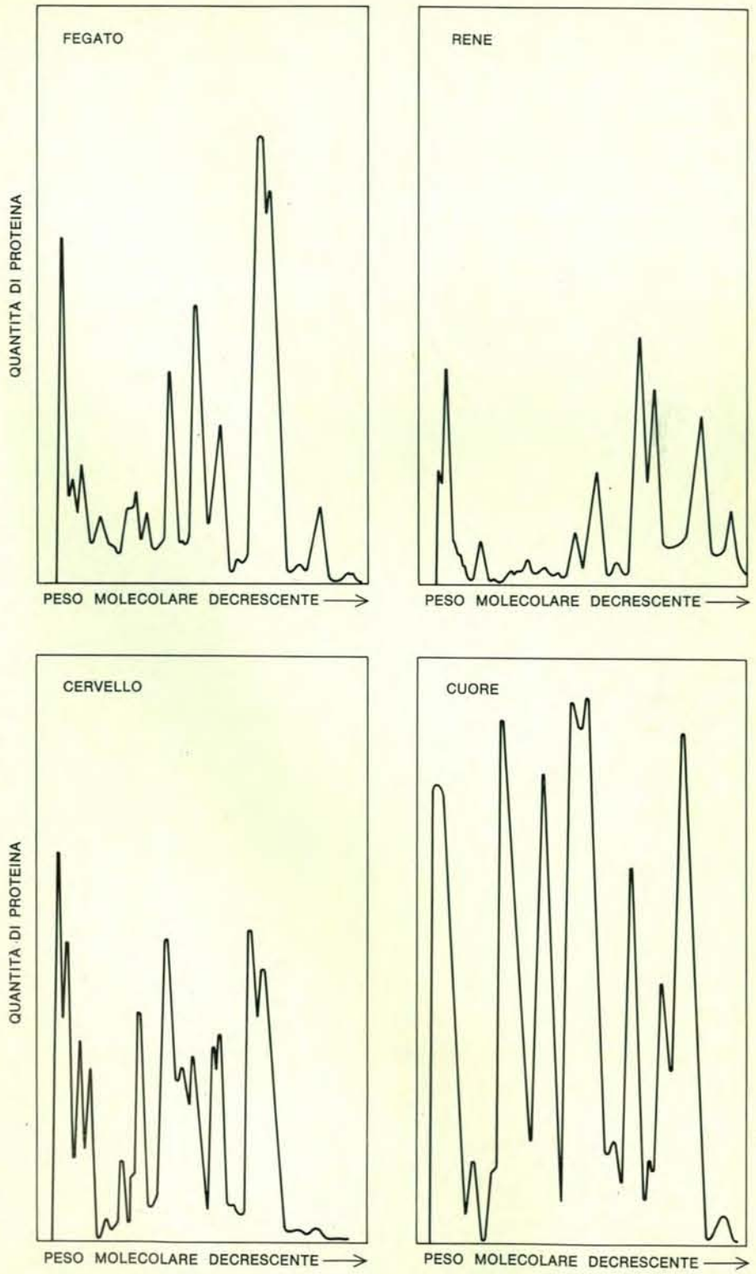
Cambiamenti nelle proteine non istoniche accompagnano anche le modificazioni che si verificano nelle proprietà strutturali e biochimiche di una cellula infettata e trasformata da virus cancerogeni. Le cellule trasformate mostrano una mancanza di controllo della crescita, che è tipica delle forme maligne. Recenti risultati ottenuti nei nostri laboratori indicano che, quasi subito dopo l'infezione delle cellule con virus tumorali, si verificano dei cambiamenti nella composizione e nel metabolismo di proteine cromosomiche non istoniche. Queste modificazioni diventano permanenti in cellule trasformate e rese neoplastiche. Finora non si è avuta alcuna prova definitiva di un rapporto causa/effetto, ma sembra logico concludere che le proteine non istoniche abbiano una funzione primaria nel mediare le modificazioni indotte da virus nell'espressione genica associate con la condizione neoplastica.

Un'altra proprietà distinta delle proteine cromosomiche non istoniche, che dà conferma all'idea che esse siano implicate nella regolazione genica specifica, è la loro capacità di legarsi al DNA. Nelle cellule batteriche in cui le proteine regolatrici specifiche sono state identificate e isolate, è risultato che esse funzionano legandosi a sequenze specifiche in regioni opportune del DNA. In maniera analoga, ci si potrebbe attendere che le molecole regolatrici specifiche, presenti nelle cellule eucariote, si leghino a tipi particolari di DNA. Nel 1970, esperimenti effettuati in uno dei nostri laboratori (precisamente in quello di Kleinsmith) hanno dimostrato che alcune proteine non istoniche si legano innanzitutto in maniera specifica al DNA.

Questi esperimenti hanno utilizzato la tecnica della cromatografia su cellulosa del DNA, sviluppata da Bruce M. Alberts della Princeton University. In essa il DNA viene adsorbito su una matrice di cellulosa che riempie un tubo di vetro. Le proteine in esame passano attraverso questa colonna; quelle che si legano al DNA, invece, si fissano alla colonna e possono poi essere rimosse per essere analizzate. Quando proteine non istoniche, preparate da fegato di ratto, si fanno passare attraverso una colonna del genere, ottenuta con DNA di ratto, una piccola percentuale delle proteine si fissa. Quando le proteine che così si legano al DNA vengono rimosse dalla colonna e fatte passare attraverso una nuova colonna, ottenuta con DNA batterico, la maggior parte passa oltre (*si veda la figura a pagina 23*). In altre parole, una parte delle proteine non istoniche deve essere in grado di riconoscere sequenze nucleotidiche specifiche nel DNA, per cui si lega al primo DNA che incontra, ma non a un DNA estraneo. Una simile interazione specifica tra queste proteine e il DNA è esattamente quello che ci si aspetterebbe se una parte di esse fosse implicata nella regolazione dell'attività di geni specifici. Questo legarsi specifico delle proteine non istoniche al DNA è stato provato da osservazioni effettuate in parecchi altri laboratori, tra cui quelli di Allfrey, Bonner e Lubomir S. Hnilica dell'Università del Texas a Houston.

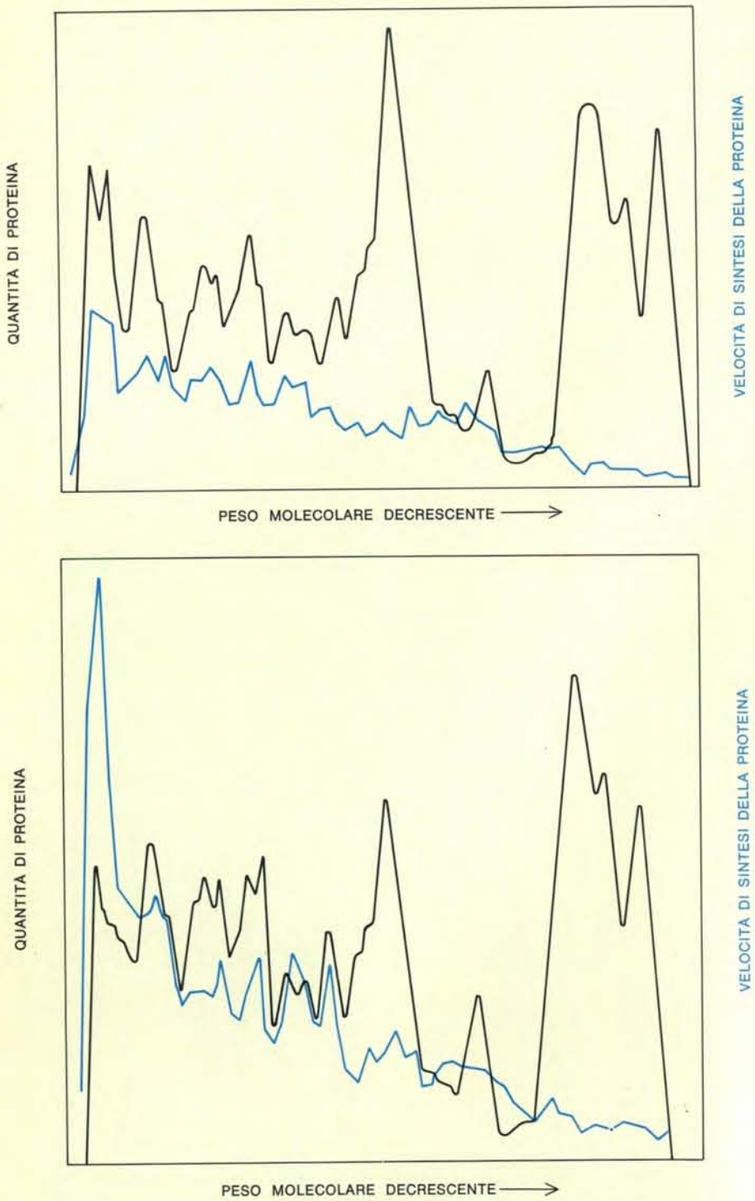
Prove sperimentali

Per determinare direttamente se la presenza delle proteine non istoniche può alterare la trascrizione genica è necessario utilizzare sistemi in cui gli effetti delle proteine non istoniche sulla sintesi dell'RNA possono essere misurati con esattezza. Nel 1968, una relazione presentata dal laboratorio di Tung-Yue Wang dell'Università dello stato di New York a Buffalo ha sostenuto che l'aggiunta di proteine non istoniche a un sistema acellulare contenente DNA e RNA polimerasi purificate può rendere reversibile l'inibizione del-



Una vasta gamma di proteine non istoniche è presente nei diversi tessuti. Ogni curva registra le fluttuazioni nella densità ottica, e quindi nella quantità delle proteine presenti, che avvengono in un gel. La disposizione dei picchi è diversa per i gel contenenti proteine non istoniche estratte da quattro diversi tessuti specializzati del ratto.

la sintesi di RNA, normalmente prodotta dalla presenza degli istoni. Questi risultati sono conformi a una precedente relazione di Thomas A. Langan, che lavorava a quell'epoca alla Rockefeller University, relazione in base alla quale quando l'istone forma un complesso con la proteina non istonica, esso esercita soltanto un parziale effetto inibitorio sulla sintesi dell'RNA. Simili modi di affrontare il problema sono stati messi in atto da Chin-Sung Teng e da Terrell H. Hamilton della Università del Texas ad Austin e da Tnomas C. Spelsberg e da Hnilica a Houston. Ambedue i gruppi hanno trovato che le proteine cromosomiche non istoniche impediscono la completa inibizione della sintesi dell'RNA da parte degli istoni. Il gruppo di Wang è andato oltre, dimostrando che la trascrizione aumentata, prodotta dall'aggiunta di proteine non istoniche, rappresenta l'attivazione di regioni del genoma precedentemente represse. Presi nel loro complesso, questi studi hanno dimostrato chiaramente che le proteine cromosomiche non istoniche possono influenzare la trascrizione. Tuttavia un'interpretazione definitiva dei risultati è parsa difficile perché i componenti della cromatina nei sistemi acellulari erano mescolati in condizioni che potevano non permettere loro di associarsi come accade nel nucleo di una cellula intatta.



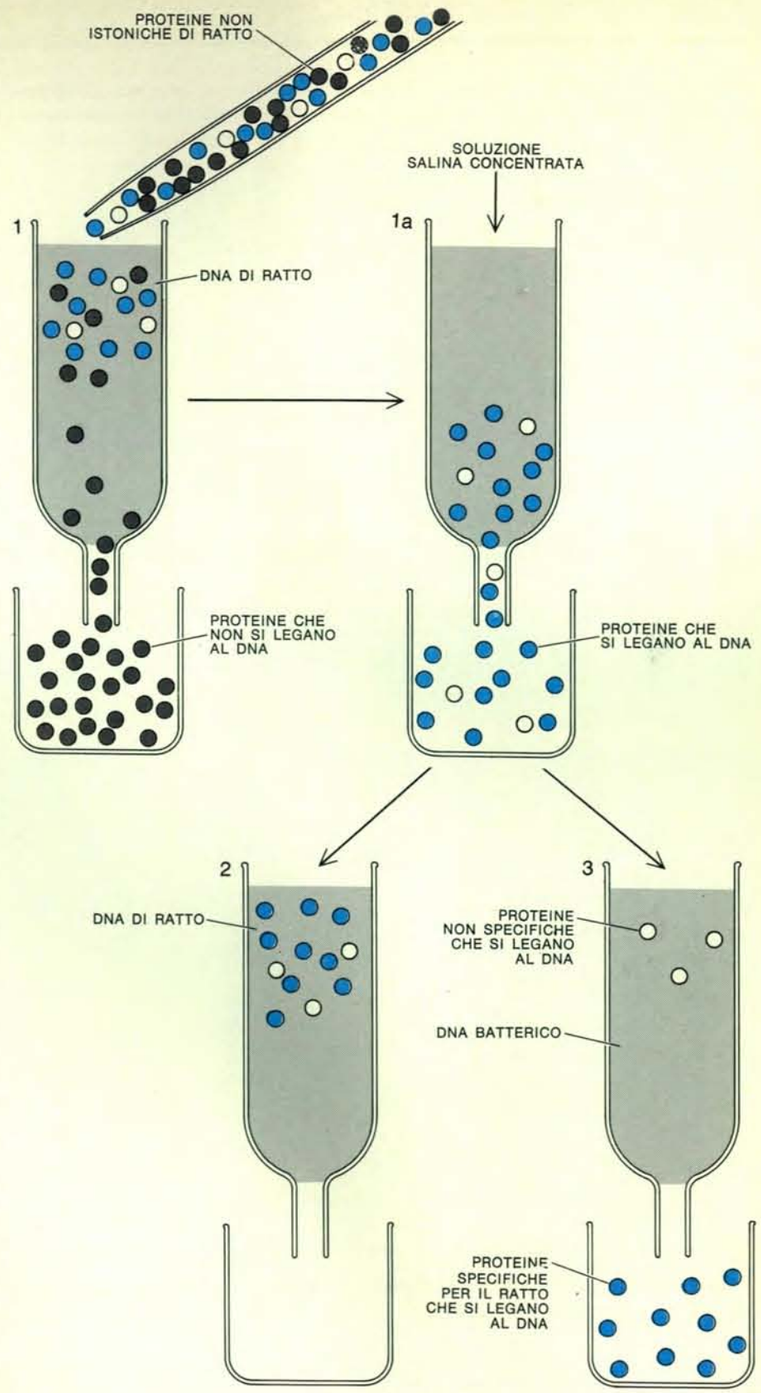
La velocità di sintesi di certe proteine non istoniche, e, di conseguenza, la quantità di ognuna di esse, varia nello stesso tipo di cellula nei diversi stadi del differenziamento. Le curve in colore mostrano le velocità di sintesi, mentre le curve in nero le quantità di proteine non istoniche che si trovano nelle cellule differenziate di un neuroblastoma di topo (*in alto*) e in altre cellule non differenziate (*in basso*). Nello stato differenziato vi è una diminuzione della velocità di sintesi delle proteine ad alto peso molecolare.

La prova più diretta che le proteine cromosomiche non istoniche svolgono una funzione nella regolazione della trascrizione specifica del tessuto proviene da esperimenti di ricostituzione della cromatina durante i quali i vari componenti di quest'ultima sono isolati e poi riuniti di nuovo in varie combinazioni. In esperimenti pionieristici, effettuati secondo questi orientamenti, R. Stewart Gilmour e John Paul del Beatson Institute for Cancer Research di Glasgow hanno isolato cromatine provenienti dal timo e dal midollo osseo di coniglio e le hanno dissociate in DNA, proteine istoniche e non istoniche. Quindi essi hanno ricostituito le cromatine ricorrendo a un metodo elaborato da Huang e Bonner. Nell'esperimento di controllo, il DNA, gli istoni e le proteine non istoniche di timo di coniglio sono stati semplicemente rimessi di nuovo insieme. La cromatina ricostituita è stata quindi utilizzata come stampo per la sintesi dell'RNA in presenza di RNA polimerasi. L'RNA trascritto si è comportato, negli esperimenti eseguiti per valutare la sua capacità di legarsi al DNA estratto da timo o da midollo osseo di coniglio, proprio come RNA di timo sintetizzato normalmente. Quindi Gilmour e Paul hanno messo insieme il DNA e gli istoni ottenuti dal timo e dal midollo osseo, e hanno ricostituito la cromatina aggiungendo al miscuglio proteine non istoniche. L'RNA prodotto da questa cromatina ibrida si è comportato esso pure come l'RNA di timo, anche se la cromatina includeva DNA e istoni di midollo osseo. D'altra parte, quando il DNA e gli istoni riuniti sono stati ricostituiti con proteine non istoniche di midollo osseo, l'RNA ottenuto dalla cromatina ibrida si è comportato proprio come l'RNA di midollo osseo (*si veda la figura a pagina 24*). Gli esperimenti hanno messo chiaramente in luce che è la presenza di proteine non istoniche, specifiche dei tessuti, a determinare quali geni saranno esattamente trascritti nei vari tessuti.

Che dire di ciò che avviene negli stessi tessuti in tempi diversi? Forse le proteine non istoniche sono responsabili anche dei cambiamenti transitori nell'attività di trascrizione, associati con i cambiamenti dello stato metabolico di un singolo tipo cellulare? Il ciclo della cellula costituisce un sistema modello ideale per analizzare in maniera diretta l'influenza che le proteine non istoniche hanno sull'attività di trascrizione del genoma, dato che la cromatina isolata dalle cellule della fase *S* (il periodo di replicazione attiva del DNA) ha una maggior attività nella sintesi dell'RNA della cromatina isolata da cellule in mitosi. Esperimenti di ricostituzione della cromatina, condotti da uno di noi (Gary Stein) e da John Farber alla Temple University, hanno dimostrato che sono le proteine non istoniche a essere responsabili di questa modulazione della trascrizione genica durante il ciclo della cellula. La cromatina, ricostituita da DNA e istoni mescolati con proteine non istoniche provenienti da cellule mitotiche, ha mostrato una minor capacità di sintesi dell'RNA della cromatina ricostituita, in presenza di proteine non istoniche provenienti da cellule durante la fase *S* (*si veda la figura a pagina 25*). Per eliminare la possibilità che gli istoni siano coinvolti in queste alterazioni gli esperimenti di ricostituzione sono stati anche effettuati in presenza di istoni provenienti da cellule in fase *S* oppure da cellule in mitosi. In questo ultimo caso, le attività di trascrizione dei due tipi di cromatine ricostituite erano identiche, mostrando così che gli istoni non stabiliscono le differenze di disponibilità delle sequenze di DNA per la sintesi dell'RNA durante il ciclo della cellula.

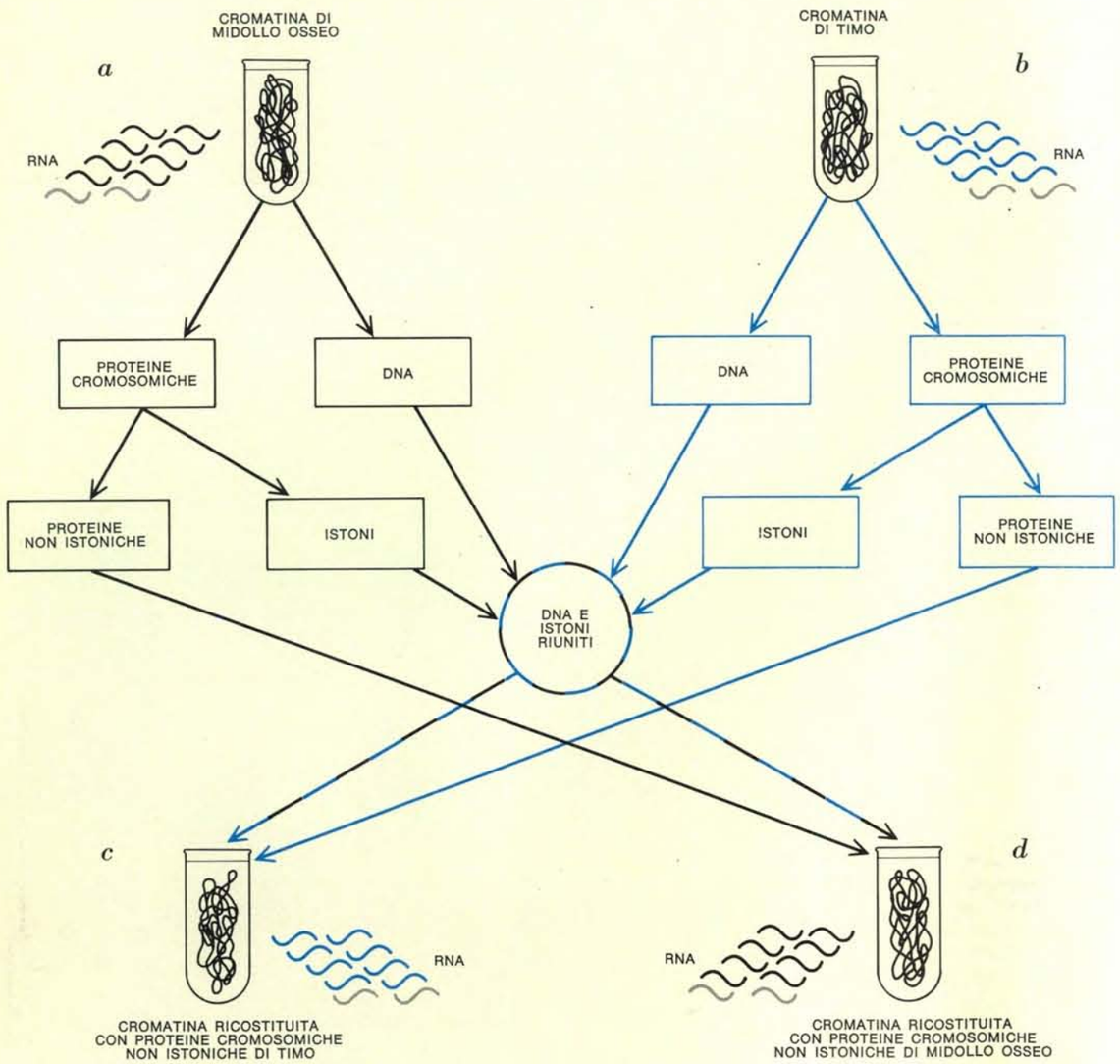
Per dimostrare chiaramente che le proteine non istoniche regolano l'attività di singoli geni, è necessario provare la trascrizione di un gene specifico. Negli ultimissimi anni, è stato possibile sintetizzare in laboratorio copie radioattive di singole sequenze geniche, risalendo da un RNA noto al suo DNA complementare. Queste copie di geni radioattivi si legano in maniera specifica agli RNA complementari per i quali codificano, e così possono essere utilizzate come «sonde» sensibili per mettere in evidenza la presenza di RNA trascritti da geni particolari (*si veda la figura in alto a pagina 26*).

Gilmour e Paul hanno utilizzato di recente queste «sonde» sofisticate in una serie di esperimenti di ricostituzione della cromatina. Questa è stata



L'associazione specifica di proteine non istoniche al DNA è stata dimostrata facendo passare proteine non istoniche di fegato di ratto attraverso una colonna di cellulosa, sulla quale era stato adsorbito del DNA di ratto. Una parte delle proteine rimaneva sulla colonna, legata al DNA; le proteine non legate (*grigio scuro*) fuoriuscivano invece dalla colonna (*1*). Le proteine legate (*in colore e in bianco*) sono state poi fatte migrare da una soluzione salina concentrata (*1a*). Esse sono state fatte passare attraverso colonne contenenti DNA di ratto o DNA del batterio *Escherichia coli*. Tutte si sono legate al DNA di ratto (*2*), mentre soltanto una piccola percentuale si è invece legata al DNA di *Escherichia coli* (*3*). Le proteine che si sono legate soltanto al DNA di ratto, ma non al DNA di *Escherichia coli* erano chiaramente specifiche per il DNA di ratto.

isolata da fegato di feto di topo, che sintetizza normalmente la globina (la proteina dell'emoglobina) e da cervello di ratto, che invece non la sintetizza. Ogni preparato di cromatina è stato trascritto in un sistema acellulare e gli RNA sintetizzati sono stati caratterizzati saggiando la loro capacità di legarsi con le copie radioattive dei geni per la globina. Conservando lo stesso tipo di trascrizione osservato nelle cellule intatte, la cromatina di fegato fetale ha sintetizzato RNA che codificava per la globina, mentre la cromatina proveniente dal cervello non lo ha fatto. Gilmour e Paul hanno proseguito quindi i loro esperimenti dissociando le cromatine e ricostituendole in presenza di vari tipi di proteine non istoniche. Essi hanno trovato che, quando la cromatina di cervello è stata ricostituita in presenza di proteine non istoniche ottenute da tessuto cerebrale, non è riuscita a sintetizzare l'RNA per la globina, mentre quando la ricostituzione si è realizzata in presenza di proteine non istoniche provenienti da fegato fetale, il gene per la globina è stato trascritto. Simili risultati danno un valido sostegno all'idea secondo la quale è la presenza di proteine non istoniche specifiche a determinare che il gene per la globina sia trascritto in tessuto epatico fetale e non in tessuto cerebrale.

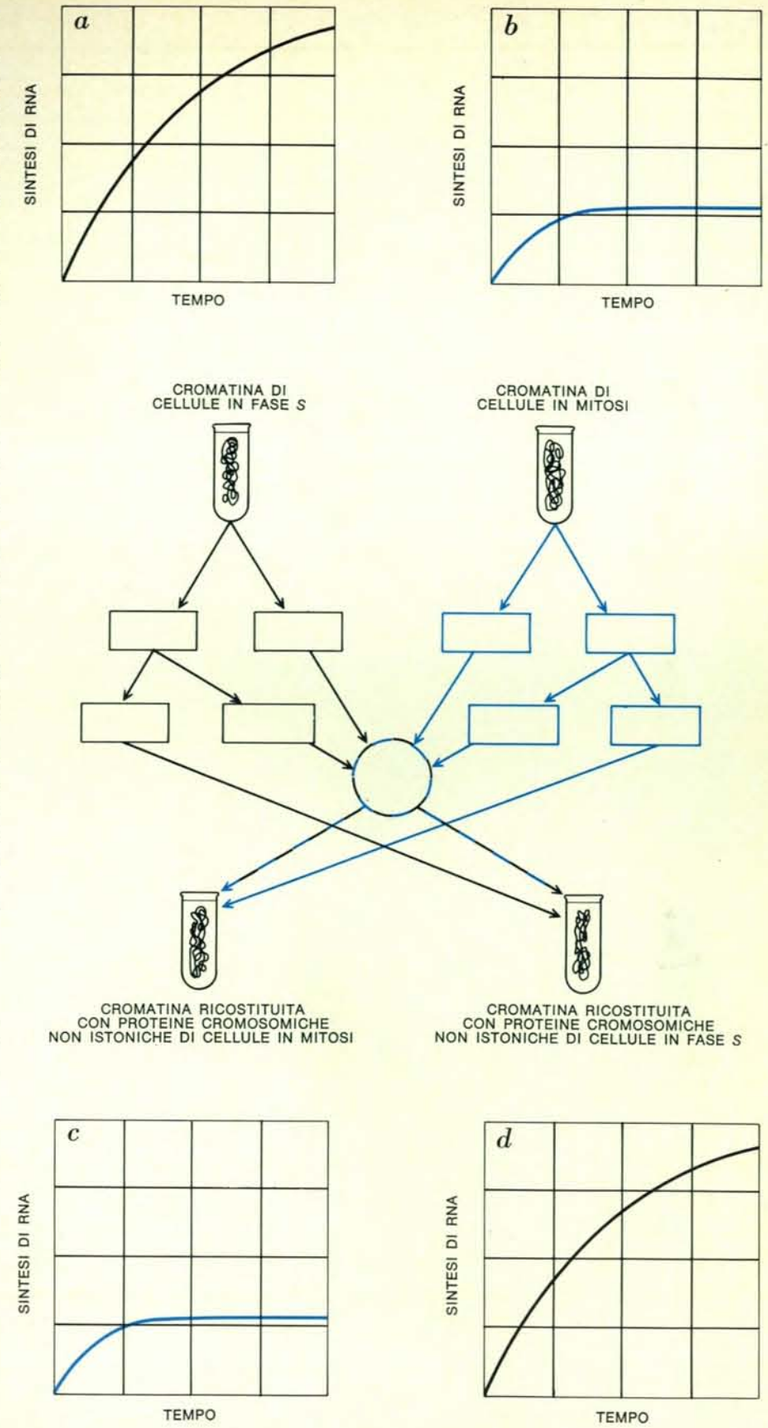


Un esperimento di ricostituzione della cromatina, effettuato da R. Stewart Gilmour e da John Paul del Beatson Institute for Cancer Research ha dimostrato che le proteine non istoniche sono responsabili della trascrizione specifica del tessuto. La cromatina è stata isolata da timo e da midollo osseo di coniglio. Ciascun tipo di cromatina ha prodotto un caratteristico RNA (*a*, *b*). La cromatina, dissociata in DNA e nelle frazioni istonica e non istonica e quindi ricostituita, ha prodotto sempre lo stesso RNA. In seguito le due cromatine sono state dissociate e sia i DNA, sia gli istoni delle due preparazioni sono stati mescolati. Quando poi il DNA e gli istoni riassociati sono stati combinati con le proteine non istoniche del timo, la cromatina ricostituita ha prodotto RNA paragonabile all'RNA di timo (*c*). Quando, invece, DNA e istoni riassociati si sono combinati con proteine non istoniche provenienti da midollo osseo, la cromatina ricostituita ha prodotto RNA paragonabile all'RNA di midollo osseo (*d*).

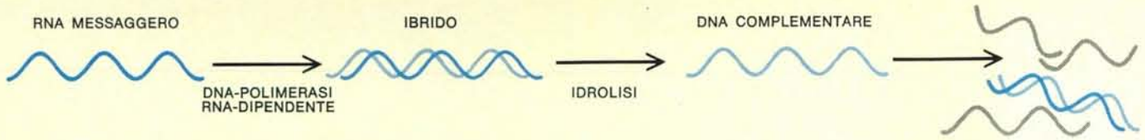
Forte di un simile modo di affrontare il problema, il nostro gruppo della Florida ha dimostrato di recente che le proteine cromosomiche non istoniche sono responsabili della limitazione della trascrizione dei geni per gli istoni alla fase *S* del ciclo cellulare (*si veda la figura in basso nella pagina successiva*). Servendosi di una «sonda» costituita da istone-DNA radioattivo per esaminare l'RNA prodotto in provetta, abbiamo dimostrato innanzitutto che la cromatina proveniente da cellule in fase *S* trascriveva gli RNA per gli istoni, mentre la cromatina ottenuta da cellule nella fase G_1 del ciclo cellulare – cioè nel periodo precedente alla replicazione del DNA – non eseguiva la trascrizione. Questo risultato concordava con una precedente dimostrazione dell'appaiamento della sintesi degli istoni con la replicazione del DNA. Abbiamo dimostrato allora che la cromatina ricostituita con proteine non istoniche da cellule in fase *S* trascriveva gli RNA per gli istoni nella stessa misura della cromatina nativa da cellule in fase *S*. Per contro, la cromatina ricostituita con proteine non istoniche isolate dalla cromatina di cellule in fase G_1 non ha mostrato nessuna trascrizione identificabile degli RNA per gli istoni. Il significato di tutto questo è che le proteine cromosomiche non istoniche governano l'elaborazione dell'informazione proveniente da geni specifici durante periodi ben definiti del ciclo della cellula.

Un'ipotesi

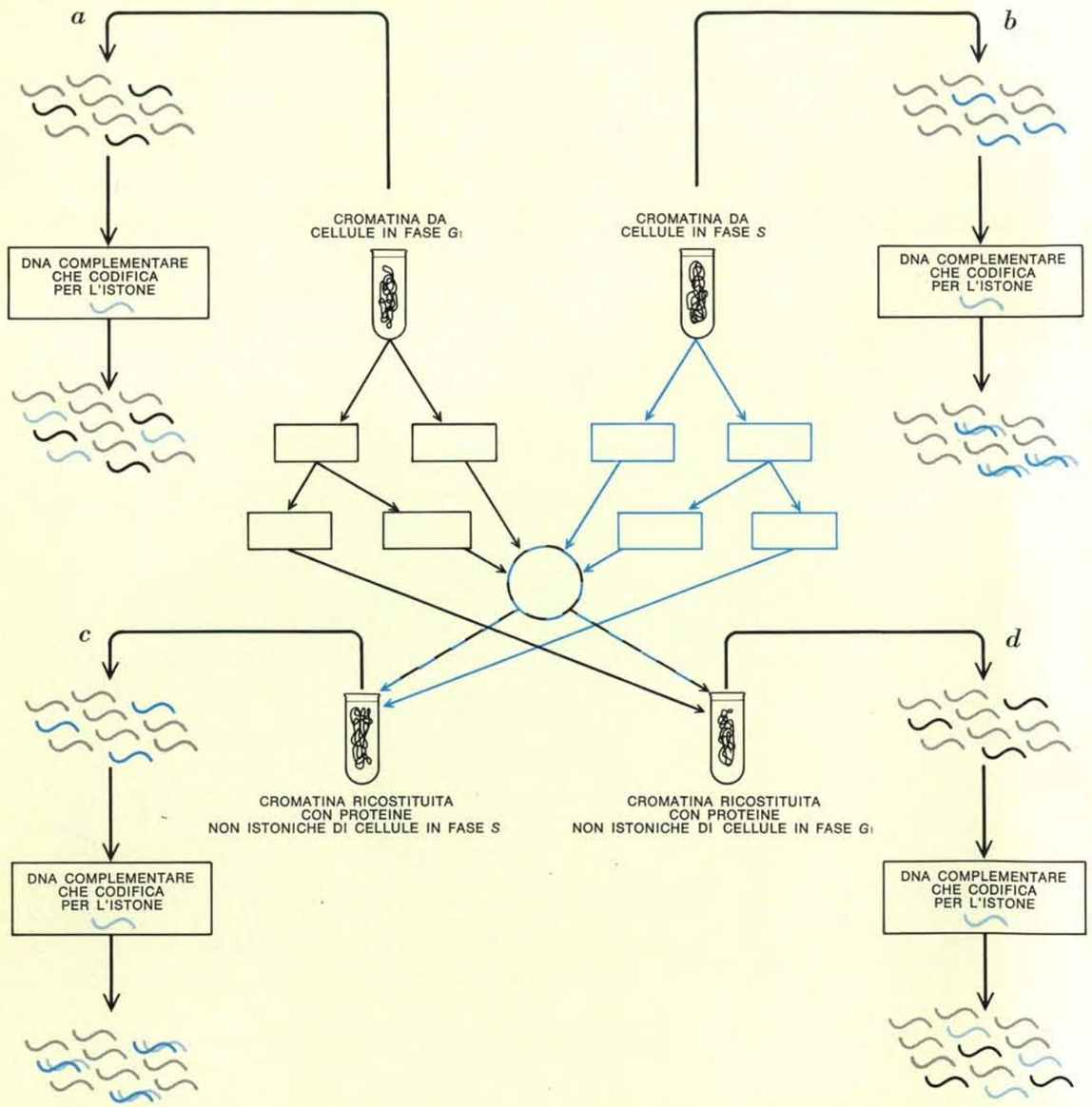
Se si ammette che le proteine cromosomiche non istoniche svolgono un ruolo chiave nel regolare la trascrizione di singoli geni sia in diverse cellule specializzate, sia in periodi differenti nello stesso tipo di cellule, ci si chiede come potrebbe essere esercitato un simile controllo. Qual è il meccanismo molecolare che ne è alla base? Una proprietà sorprendente delle proteine non istoniche, è l'ampio grado della loro fosforilazione. In una serie di esperimenti iniziati nel 1966 Kleinsmith, Allfrey e Mirsky hanno osservato che le proteine non istoniche vengono continuamente modificate nel nucleo per aggiunta e rimozione di gruppi fosfato. Abbiamo pensato che, siccome i gruppi fosfato sono dotati di carica negativa, tali modificazioni potevano alterare sia la struttura, sia le interazioni funzionali delle proteine non istoniche nella cromatina. Da quando sono state effettuate queste prime osservazioni, i ricercatori di molti laboratori,



Un altro esperimento di ricostituzione ha dimostrato che le proteine non istoniche governano le variazioni di trascrizione, che si notano in diversi stadi del ciclo della cellula. La cromatina isolata da cellule in fase *S* sintetizza più rapidamente RNA (*a*) della cromatina isolata durante la mitosi (*b*). Le due cromatine sono state dissociate e il DNA e gli istoni riuniti. La cromatina ricostituita con proteine non istoniche di cellule mitotiche ha sintetizzato RNA a una velocità caratteristica per la cromatina di cellule in mitosi (*c*). La cromatina ricostituita con proteine non istoniche di cellule in fase *S* ha prodotto dell'RNA alla maggiore velocità tipica della cromatina di cellule in fase *S* (*d*).



Gli esperimenti di ibridazione servono a identificare prodotti specifici del gene. Un RNA noto (*in colore pieno*) viene isolato. In presenza dell'enzima DNA polimerasi RNA dipendente (transcriptasi inversa), l'RNA serve da stampo per il montaggio di un filamento complementare di DNA (*in colore chiaro*), che può essere marcato mediante incorporazione di una subunità radioattiva. L'ibrido RNA-DNA viene dissociato mediante idrolisi e dà luogo a un DNA complementare radioattivo, ossia a un gene sintetizzato. Questo è utile per identificare l'RNA da cui è stato prodotto in quanto forma ibridi solo con quell'RNA (*a destra*).



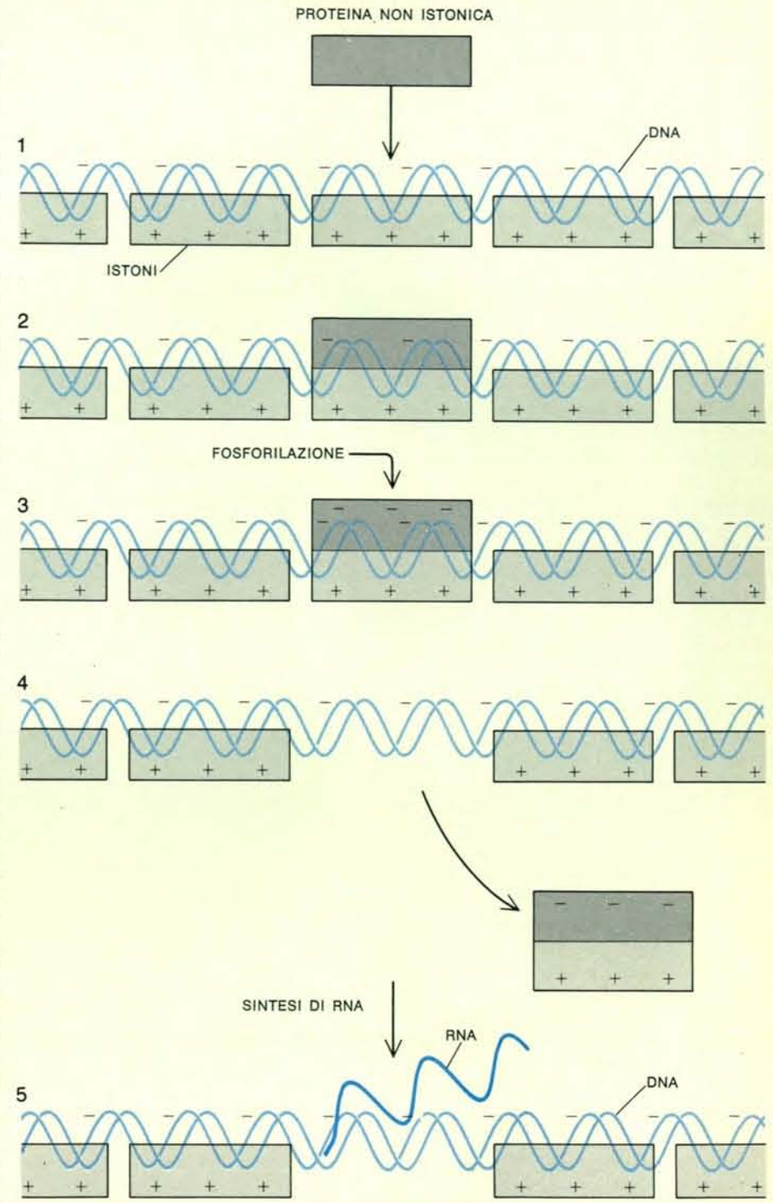
Nell'esperimento qui descritto è stato preparato un DNA radioattivo complementare all'RNA per l'istone ed è stato quindi usato come «sonda» (*in colore chiaro*), con la quale dimostrare che, mentre la cromatina proveniente da cellule in fase G_1 non produce RNA per l'istone (*a*), la cromatina di cellule in fase *S* lo fa (*in colore pieno, b*). Le cromatine sono state quindi dissociate, gli istoni e il DNA riuniti di nuovo e quindi ricostituiti con due tipi di proteine non istoniche. Gli RNA sono stati preparati da due cromatine ricostituite e quindi saggiati con il DNA. La cromatina, ricostituita con le proteine non istoniche provenienti da cellule in fase *S*, ha sintetizzato l'RNA per l'istone proprio come la cromatina naturale proveniente da cellule in fase *S* (*c*). La cromatina, ricostituita con le proteine non istoniche di cellule in fase G_1 (*d*), non è invece in grado di sintetizzare RNA.

studiando diversi sistemi biologici, hanno notato che la fosforilazione delle proteine non istoniche è associata con il differenziamento cellulare e con l'attivazione dei geni. Inoltre è stato dimostrato che la capacità delle proteine non istoniche di stimolare la sintesi dell'RNA in sistemi acellulari dipende dal loro stato di fosforilazione.

Tali osservazioni suggeriscono che la fosforilazione delle proteine non istoniche possa essere in qualche modo implicata nel meccanismo grazie al quale queste molecole regolano la trascrizione genica. Un modello proposto chiama in causa sia il legame con DNA specifico sia la fosforilazione (*si veda la figura in questa pagina*). Noi suggeriamo che una proteina non istonica specifica si leghi a un sito specifico del DNA, represso dall'istone. In quel punto la proteina non istonica diventa fosforilata e, dato che i gruppi fosfato sono dotati di carica negativa, essa comincia a respingere il DNA, che è pure dotato di carica negativa, mentre d'altra parte si associa intensamente con gli istoni dotati di carica positiva. L'effetto di questi due processi è quello di provocare il distacco del complesso proteina istonica-proteina non istonica dal DNA. La risultante area priva di DNA, mancando del suo istone inibitore, può quindi essere trascritta in RNA.

Anche se questo modello del distacco degli istoni è molto ipotetico, un certo numero di osservazioni concordano con le sue caratteristiche generali. Come abbiamo ricordato, i geni sono trascritti più attivamente durante la fase *S* che non durante la mitosi. Abbiamo osservato di recente che anche le proteine non istoniche sono più attivamente fosforilate durante la fase *S* che non durante la mitosi e inoltre che gli istoni si legano meno saldamente al DNA durante la fase *S* che non durante la mitosi. Alcuni recenti esperimenti sulla ricostituzione della cromatina, effettuati nel nostro laboratorio in Florida, hanno dimostrato che le proteine non istoniche associate con il genoma durante la fase *S* sono responsabili in maniera specifica del legame meno forte degli istoni con il DNA durante quel periodo del ciclo della cellula. Questi risultati dimostrano se non altro almeno uno stretto rapporto tra attività genica, fosforilazione della proteina non istonica e associazione degli istoni al DNA.

Il modello del distacco degli istoni è solo uno dei molti proposti per spiegare il comportamento delle proteine cromosomiche nella regolazione genica. Anche se il meccanismo particolareggiato non è stato ancora chiaramente dimostrato, la comprensione dell'attività delle proteine cromosomiche e delle loro interazioni con il DNA sta avanzando rapidamente. Certo le proteine istoniche e non istoniche svolgono delle funzioni importanti nella determinazione delle proprietà strutturali e funzionali del genoma. Mentre gli istoni sono implicati nel mantenimento della struttura della cromatina e nella repressione non specifica delle sequenze geniche, sembra chiaramente dimostrato che le proteine cromosomiche non istoniche riconoscano loci genici ben definiti e quindi regolino la trascrizione dell'informazione genetica specifica.



Il gene viene attivato, secondo un ipotetico modello, quando una proteina non istonica riconosce (*1*) un sito specifico su un filamento di DNA che generalmente viene represso da istoni. La proteina non istonica si lega al sito (*2*); quindi subisce una fosforilazione e diventa pertanto dotata di carica negativa (*3*). Poi respinge il DNA dotato di carica negativa e si allontana contemporaneamente all'istone dotato di carica positiva, lasciando nudo il sito posto sul DNA (*4*). Questo sito, che in questo modo non è più represso dall'istone, diviene disponibile per la trascrizione dell'RNA (*in colore, 5*).

Modelli a risonanza duale delle particelle elementari

In questo nuovo modello teorico le particelle a interazione forte (adroni) vengono descritte matematicamente come corde senza massa i cui estremi si muovono con la velocità della luce in uno spazio multidimensionale

di John H. Schwarz

Negli ultimi 20 anni sono state prodotte e identificate in collisioni provocate da acceleratori di particelle ad alta energia centinaia di differenti particelle «elementari». Nello stesso periodo sono stati proposti molti schemi empirici e teorici nel tentativo di mettere ordine in una massa di dati che a prima vista appare completamente caotica. Sono stati fatti notevoli progressi e molti aspetti del problema sono oggi ben conosciuti, ma la fisica appare ancora ben lontana dal raggiungimento di una teoria completa delle particelle elementari. Questo articolo espone un nuovo promettente metodo, noto come teoria della risonanza duale, che può rivelarsi utile nella comprensione del comportamento di una delle due grandi classi di particelle elementari, gli adroni, cioè le particelle soggette a una delle quattro forze conosciute della fisica: l'interazione forte.

Tutte le trattazioni teoriche della fisica delle particelle elementari sono basate su due principi solidamente affermati: la teoria della relatività ristretta e la teoria dei quanti. La teoria della relatività ristretta, oltre a enunciare la famosa relazione tra energia e massa ($E = mc^2$), descrive la correlazione esistente tra le osservazioni eseguite da osservatori in moto con velocità differenti. Le teorie sono solitamente formulate in modo covariante, il che significa che tali correlazioni sono implicite nel formalismo matematico che esprime la teoria. Secondo la teoria dei quanti, una particolare interazione sperimentale – per esempio la collisione di due particelle – può verificarsi in molti modi diversi, ciascuno con una ben determinata probabilità. Queste probabilità vengono espresse mediante numeri complessi chiamati ampiezze di diffusione. Il compito di ogni teoria delle particelle è di permettere il calcolo dell'ampiezza di diffusione di una data interazione tra particelle con un alto grado di precisione. Se si cerca di calcolare le ampiezze di diffusione (date in funzione di un adeguato insieme di variabili), anche la teoria dei quanti rientra automaticamente nel nuovo schema teorico.

Le particelle elementari sono soggette a parecchi differenti tipi di forze (o interazioni), che si possono raggruppare in due vaste categorie: forti e non forti. La categoria di quelle non forti viene di solito ulteriormente suddivisa in forza elettromagnetica, forza gravitazionale e forza debole. Per gli scopi di questa trattazione possiamo ignorare quest'ultima distinzione, poiché le interazioni forti sono l'oggetto più importante della teoria della risonanza duale. Le interazioni forti operano tra le particelle elementari solo a brevi distanze (tipicamente dell'ordine di 10^{-13} centimetri) e sono circa 100 volte più intense delle più forti tra le altre interazioni, cioè quelle elettromagnetiche. Le interazioni forti sono quelle che tengono insieme il nucleo atomico (per cui sono dette anche nucleari) e sono in gran parte responsabili delle proprietà dei protoni e dei neutroni che costituiscono il nucleo.

Dal punto di vista filosofico, sono propenso a credere che dovrebbe essere possibile trovare una descrizione matematica di tutte le particelle e delle loro interazioni che sia completa, consistente e persino elegante. Per progredire verso la formulazione di una tale teoria, sembra opportuno affrontare dapprima problemi più modesti mediante approssimazioni successive. Due sistemi ipotetici più semplici da esaminare sono: (1), adroni soggetti solo a interazioni forti e, (2), non adroni soggetti solo a interazioni non forti. Ciascuno dei due sistemi è nel migliore dei casi solo un'approssimazione, poiché in effetti gli adroni sono soggetti anche alle interazioni non forti. È immaginabile che entro ogni sistema si possa trovare una teoria autoconsistente che possa descrivere con precisione vaste classi di esperimenti, ma è purtroppo altrettanto probabile che ciascuna delle suddette approssimazioni implichi l'eliminazione di elementi logici che sono essenziali per delineare una teoria matematicamente consistente di tutte le particelle elementari. Per il momento possiamo solo sperare che si avveri la prima possibilità. In conseguenza di questa ipotesi il resto della trattazione verterà sugli adroni e sulle loro interazioni forti, con l'ipotesi sottintesa che si possa successivamente trovare una teoria consistente e completa di tutte le interazioni.

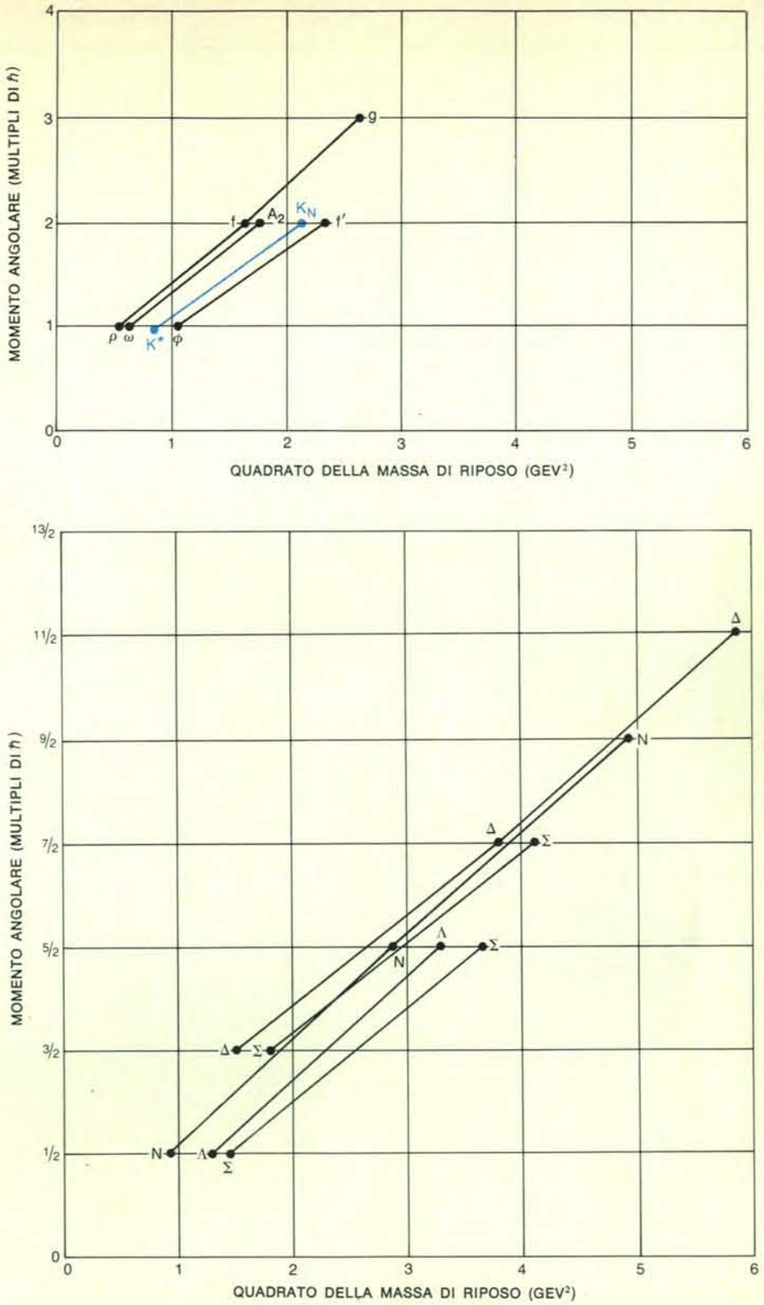
Le conoscenze teoriche sugli adroni ottenute finora sono basate principalmente su metodi che non tentano di dare formule esatte per tutte le possibili ampiezze di diffusione, ma descrivono piuttosto importanti aspetti generali degli adroni che dovrebbero essere contenuti o spiegati in una teoria completa. Esempi di questi metodi di indagine comprendono vari modelli statistici e schemi di classificazione delle particelle.

Un'interessante possibilità che deriva da uno studio di modelli statistici compiuto da Rolf Hagedorn dell'Organizzazione Europea per le Ricerche Nucleari (CERN) di Ginevra è che possa esistere un numero infinito di adroni, la cui distribuzione sia una funzione esponenziale crescente della massa. Una siffatta distribuzione implicherebbe l'esistenza di una temperatura universale massima (circa 2×10^{12} kelvin). In favore del tipo di distribuzione di Hagedorn vi è qualche indizio sperimentale, non ancora confermato. Lo stesso tipo di spettro di particelle nasce anche in modo naturale (e inevitabile) nei modelli a risonanza duale. Questo aspetto pone inoltre tali modelli in netto contrasto con i modelli della usuale teoria quantistica dei campi, che ne costituiscono i principali concorrenti nella ricerca di una teoria degli adroni. I modelli della teoria dei campi riescono solo con difficoltà a rendere conto di una tale distribuzione di massa, anche se non la escludono totalmente.

Il significato di una temperatura universale può essere valutato considerando la seguente ipotetica situazione. Supponiamo che una pentola di adroni sia prossima a tale temperatura massima e che si tenti di portarla a una temperatura superiore fornendo energia, per esempio bombardando gli adroni con particelle ad alta energia in condizioni tali da dare origine a un equilibrio termico. Le previsoni dell'ipotesi di Hagedorn sono che una volta raggiunta la temperatura massima, qualsiasi ulteriore cessione di energia si traduca nella creazione di nuove particelle piuttosto che in un aumento di temperatura. Qualcuno ha avanzato l'ipotesi che un siffatto «quarto principio della termodinamica» possa avere influenzato l'evoluzione dell'universo fino dai primi stadi del *big bang* (grande scoppio iniziale).

Vi sono due schemi distinti di classificazione delle particelle che rendono conto con notevole successo dei differenti aspetti dello spettro di particelle elementari finora osservato. Il primo (basato sulle ricerche di Murray Gell-Mann e di George Zweig dell'Istituto di Tecnologia di California) è chiamato modello a quark. Esso prevede, in ultima analisi, il raggruppamento degli adroni in certe famiglie, nelle quali i diversi componenti sono simili, fatta eccezione per certi numeri quantici, o proprietà fondamentali, quale la carica elettrica (esempi di numeri quantici di significato più riposto sono le proprietà chiamate spin isotopico e ipercarica). Il modello a quark stabilisce inoltre che il criterio di raggruppamento può apparire chiaro se si suppone che gli adroni siano formati secondo certe regole da tre costituenti fondamentali (i quark) e dalle loro antiparticelle (antiquark). Questa schematizzazione ricorda da vicino la suddivisione di molecole in atomi, di atomi in elettroni e nuclei, di nuclei in neutroni e protoni; ora si afferma che i neutroni e i protoni sono a loro volta costituiti da quark.

A questo punto le analogie paiono crollare, poiché non è ancora stata osservata nelle collisioni ad alta energia la produzione di quark come particelle libere. Per questo e altri motivi



Le traiettorie di Regge, curve ipotetiche lungo le quali possono essere raggruppate le famiglie di adroni con differenti valori dello spin e della massa, risultano essere pressoché rettilinee e approssimativamente parallele, in accordo con le previsioni della teoria della risonanza duale. Nella figura in alto sono indicate con linee oblique alcune delle traiettorie dei mesoni meglio conosciute. I simboli indicano i componenti finali di ciascuna famiglia di mesoni secondo diversi sistemi di nomenclatura; alcuni dei mesoni indicati sono risonanze estremamente instabili non elencate nella tabella della pagina seguente. Le particelle dotate della proprietà chiamata stranezza sono indicate a colori. Nella figura qui sopra sono riportate le più note traiettorie dei barioni. In questo caso *N* indica un nucleone, che può essere o un protone o un neutrone. L'unità di momento angolare, $\hbar$ è uguale alla costante di Planck (indicata con *h*) divisa per 2π. I mesoni hanno valori interi del momento angolare, mentre i barioni hanno valori semiinteri. La massa di riposo è espressa in unità equivalenti di energia secondo la relazione di Einstein $E = mc^2$; GeV indica il gigaelettronvolt, pari a un miliardo di elettronvolt. Tutte le traiettorie nelle due figure hanno una pendenza di poco inferiore a $1/\text{GeV}^2$.

i teorici credono oggi che i quark siano permanentemente legati all'interno degli adroni e non possano esser prodotti in nessun caso come particelle libere. Un certo numero di tentativi per comprendere il confinamento permanente dei quark entro gli adroni è stato fatto nel contesto sia dei modelli della teoria dei campi sia dei modelli a risonanza duale, ma finora nessuna delle proposte avanzate è molto convincente. Al punto in cui stanno le cose ora, il problema del « confinamento dei quark » rappresenta un serio e forse fondamentale ostacolo a tutti gli sforzi fin qui compiuti per costruire una teoria completa delle particelle elementari.

Il secondo importante schema di classificazione delle particelle non ha un nome semplice, anche se ha avuto molto successo. Secondo tale schema gli adroni possono essere raggruppati in famiglie i cui costituenti hanno valori diversi del momento angolare e della massa, ma sono simili sotto tutti gli altri aspetti (il momento angolare, secondo la teoria dei quanti, può avere solo valori discreti, multipli della costante di Planck ridotta). I componenti di ciascuna famiglia giacciono su una ipotetica curva detta traiettoria di Regge (dal nome del fisico torinese Tullio Regge). L'idea che tutti gli adroni giacciano su tali traiettorie, chiamata talvolta ipotesi di Regge (anche se il lavoro originale di Regge non riguardava gli adroni), è oggi largamente accettata.

ADRONI	MESONI	PIONI (π^+, π^0, π^-)
		KAONI ($K^+, K^0, \bar{K}^0, K^-$)
		ETA (η^0)
		RHO (ρ^+, ρ^0, ρ^-)
		OMEGA (ω)
	BARIONI	PROTONE (p)
		NEUTRONE (n)
		LAMBDA (Λ^0)
		SIGMA ($\Sigma^+, \Sigma^0, \Sigma^-$)
		DELTA ($\Delta^{++}, \Delta^+, \Delta^0, \Delta^-$)
NON ADRONI	LEPTONI	ELETTRONE (e^-)
		POSITONE (e^+)
		MUONI (μ^+, μ^-)
		NEUTRINI ($\nu_e, \bar{\nu}_e, \nu_\mu, \bar{\nu}_\mu$)
	ALTRI	FOTONE (γ)
		GRAVITONE (NON ANCORA OSSERVATO)

Gli adroni e i non adroni, rispettivamente particelle soggette a interazioni forti e particelle che non lo sono, sono ulteriormente suddivisi in quattro sottoclassi in questo elenco parziale che comprende, tra le particelle elementari osservate, solo quelle a vita media relativamente più lunga e meno pesanti. I mesoni e i barioni, le due sottoclassi degli adroni, sono contraddistinti da vari numeri quantici, come il momento angolare intrinseco, o spin. I simboli soprassegnati indicano un'antiparticella; in alcuni casi la particella coincide con la propria antiparticella. I neutrini e gli antineutrini possono essere di due tipi: elettronico, ν_e e muonico, ν_μ. Negli acceleratori di particelle sono stati prodotti centinaia di adroni più pesanti a vita media più breve (chiamati anche risonanze).

I fatti empirici sono molto più specifici e vanno persino al di là di quanto suggerisce l'ipotesi di Regge. Le traiettorie di Regge risultano sperimentalmente quasi rettilinee, il che significa che il momento angolare delle particelle su una particolare traiettoria è espresso con buona approssimazione da una funzione lineare del quadrato della massa della particella (*si veda la figura nella pagina precedente*). Inoltre, tutte le traiettorie di Regge per gli adroni consueti appaiono approssimativamente parallele. Ulteriori prove sperimentali a sostegno dell'ipotesi di Regge e della linearità delle ordinarie traiettorie di Regge si ricavano dall'osservazione del comportamento alle alte energie delle funzioni di probabilità di interazione note ai fisici delle particelle come sezioni d'urto di diffusione. I modelli a risonanza duale godono della proprietà che nella prima approssimazione tutte le particelle giacciono su traiettorie di Regge parallele e rettilinee. Questo fatto assicura che lo spettro delle particelle e le ampiezze di diffusione hanno molti aspetti realistici. D'altra parte, nei modelli della teoria dei campi è difficile render conto della linearità delle traiettorie di Regge. A mio avviso questa proprietà dei modelli a risonanza duale è una delle ragioni più stringenti per affermare la validità di tale linea d'attacco.

In pratica la teoria della risonanza duale è lo studio di una classe di modelli matematici. Il principale obiettivo è ovviamente quello di trovare l'esempio particolare che si presti a descrivere il nostro universo, ma i modelli finora realizzati non sono evidentemente in grado di farlo, anche se presentano molti aspetti che fanno ritenere di essere sulla strada giusta, quali per esempio la distribuzione di Hagedorn e le traiettorie rettilinee parallele di Regge. Un'altra importante proprietà di questi modelli è implicita nel loro nome. Cosa si intende per risonanza duale?

Il termine « dualità » è stato usato nella fisica delle alte energie con parecchi significati differenti. Per evitare confusione possiamo definire separatamente la « dualità teorica » dalla « dualità fenomenologica ». La dualità fenomenologica, dal punto di vista storico, è stata scoperta per prima (nel 1967) e ha avuto un ruolo importante nel motivare la formulazione della dualità teorica un anno dopo. La dualità fenomenologica è il fatto empirico che quando il regolare comportamento asintotico di una sezione d'urto di diffusione ad alta energia viene estrapolata all'indietro verso le energie inferiori, la curva che ne risulta tende a rappresentare una media dei massimi e minimi della vera sezione d'urto (*si veda la figura nella pagina a fianco in alto*). I massimi sono dovuti a particelle instabili a breve vita media chiamate risonanze. Anche se non manca qualche ambiguità nel procedimento di estrapolazione, la dualità fenomenologica ci permette di ricavare relazioni approssimate tra i parametri che descrivono il comportamento alle alte energie e le masse e le intensità di accoppiamento delle risonanze a energie inferiori.

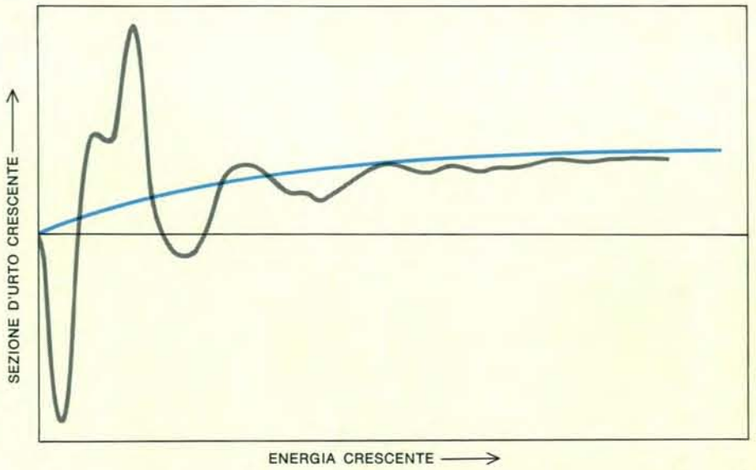
La dualità teorica è una proprietà intrinseca nella costruzione dei modelli a risonanza duale. Si tratta di una precisa affermazione matematica sulla struttura delle ampiezze di diffusione calcolate in prima approssimazione, nella quale le formule sono relativamente semplici poiché le risonanze non hanno larghezza (cioè non presentano dispersione in energia). I modelli a risonanza duale vengono calcolati con una serie di approssimazioni successive che partono e sono determinate da questa prima approssimazione. L'intera serie costituisce la teoria completa, ma la dualità teorica riguarda solo la prima approssimazione. L'esatto enunciato della dualità teorica è che l'ampiezza di diffusione calcolata sommando tutti i contributi degli « scambi » di risonanze uguaglia quella ottenuta sommando tutte le risonanze di « canale diretto » (*si vedano le figure nella pagina a fianco in basso*). Questa formulazione rassomiglia molto all'ipotesi del *bootstrap* (stringa da scarpe) avanzata da Geoffrey Chew dell'Università di California di Berkeley e da Steven Frautschi dell'Istituto di Tecnologia della California. Secondo la loro ipotesi, lo scambio di risonanze fornisce le forze necessarie per tenere insieme le risonanze di canale diretto. Una differenza è che il meccanismo di dualità non determina l'intensità complessiva della forze. Il nuovo aspetto dei modelli a risonanza duale che rende possibile questo fatto è che essi sono basati su un'infinità di risonanze con proprietà strettamente correlate la cui funzione specifica è di incorporare in prima approssimazione il meccanismo di bootstrap.

Dal punto di vista storico i modelli a risonanza duale vennero sviluppati inventando formule per le ampiezze di diffusione nella prima approssimazione. Il processo fu iniziato da Gabriele Veneziano dell'Istituto Scientifico Weizmann (Israele) e del CERN e fu continuato da molti altri. La teoria venne poi studiata allo scopo di dimostrare che essa soddisfa certi principi generali, quali per esempio i requisiti che le probabilità di interazione siano positive e diano per somma 1 e che un segnale non possa essere ricevuto prima ancora di essere stato inviato. Negli ultimi anni è stato intrapreso con successo un programma di ricerca che spiegasse come queste e altre proprietà siano intrinseche alla teoria della risonanza duale. Questo metodo di attacco ha avuto uno stimolante impulso dal recente lavoro di Yoichiro Nambu dell'Università di Chicago, Leonard Susskind dell'Università di Yeshiva (New York), Holger Nielsen dell'Istituto Niels Bohr di Copenaghen e altri, i quali hanno mostrato l'esistenza di un semplice quadro fisico degli adroni previsto dai modelli duali. Questo quadro fornisce un modo elegante per la descrizione della teoria della risonanza duale, anche se non è sempre il più conveniente per i calcoli dal punto di vista pratico.

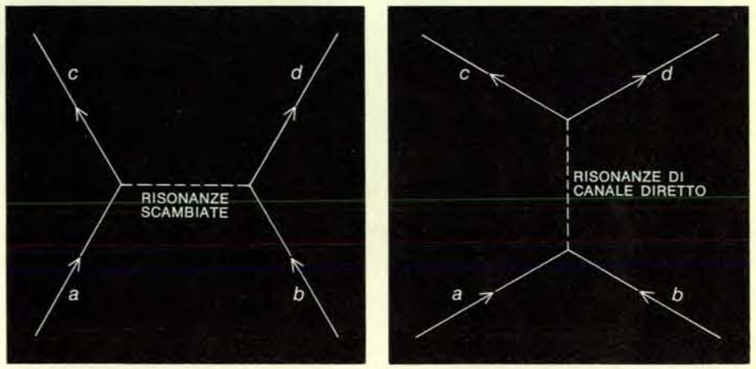
Nambu e gli altri hanno scoperto che in un modello di risonanza duale gli adroni si possono considerare come stati quantici di una corda, anche se di un tipo molto particolare di corda. Da un lato la corda è intrinsecamente priva di massa e quindi il suo moto è necessariamente relativistico; perciò, per esempio, i suoi estremi devono muoversi sempre con la velocità della luce. Corde con queste proprietà sono dette corde di luce. La massa di un adrone nasce dall'energia dovuta alla tensione della corda e dall'energia cinetica del suo moto rispetto al centro di massa (la tensione, ricavata dalla pendenza delle traiettorie di Regge, risulta essere di circa 13 tonnellate). Emerge da questo fatto la differenza tra i modelli di risonanza duale e i modelli della teoria dei campi: le particelle elementari descritte secondo un modello della teoria dei campi sono puntiformi, mentre quelle di un modello di risonanza duale hanno un'estensione spaziale.

Una volta accettata la possibile esistenza di un'estensione spaziale è naturale chiedersi perché mai la struttura di un adrone debba essere monodimensionale (a corda) piuttosto che bi- o tridimensionale (a goccia). Per definizione i modelli a risonanza duale accettano solo una struttura monodimensionale. È stata esaminata anche la possibilità di una struttura a gocce, ma tutto ciò che sappiamo finora è che queste teorie sono molto più difficili dal punto di vista matematico. Secondo me, gli schemi esistenti basati sulle corde sono sufficientemente realistici da rendere plausibile che una corretta teoria degli adroni possa essere basata su un tipo di corda ancor più fantastico di tutti quelli fin qui ipotizzati.

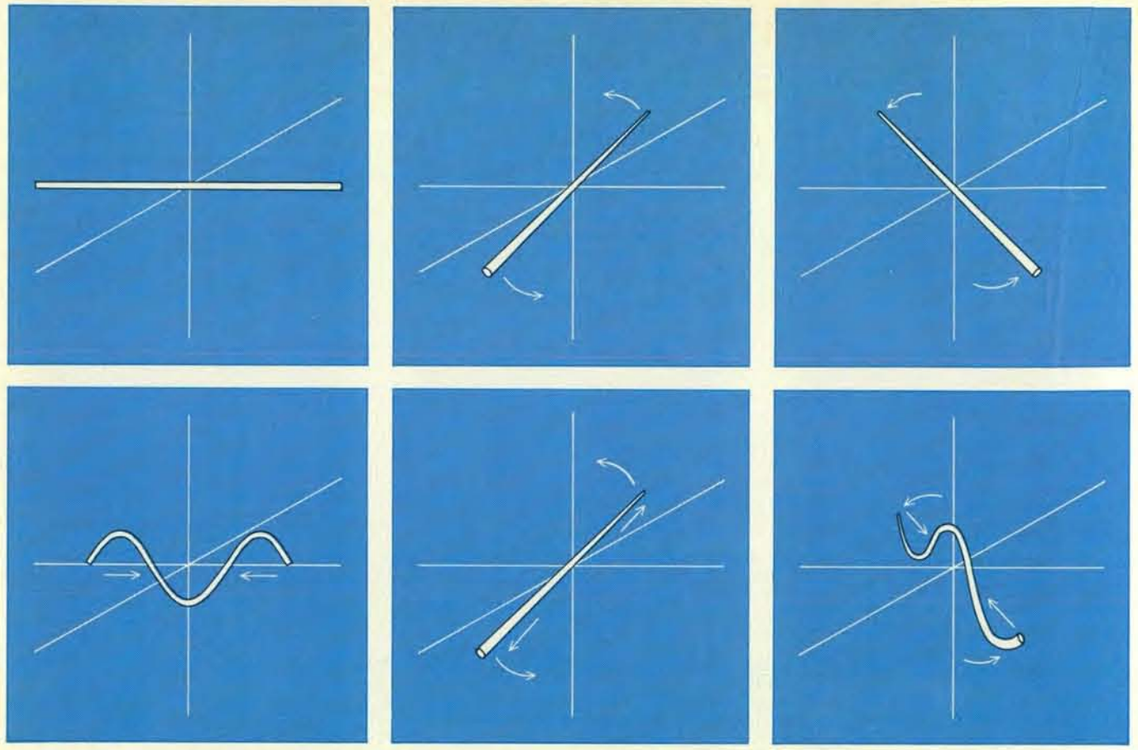
La teoria delle corde di luce avrebbe potuto essere studiata 40 anni fa, ma non vi era alcun motivo per farlo fin quando non si sospettò una correlazione con i modelli duali. Questo studio è un'applicazione interessante e istruttiva dei principi della relatività ristretta e della meccanica quantistica. Se si esaminano solo i moti classici (cioè non quantistici) di corde non interagenti, il problema non presenta alcuna difficoltà, ma se si inseriscono le imposizioni della teoria dei quanti gli effetti sono notevoli. Una sorpren-



La dualità delle risonanze è definita empiricamente mediante l'osservazione del comportamento delle funzioni di probabilità di interazione chiamate sezioni d'urto di diffusione. In questo esempio la curva in grigio rappresenta la differenza tra la sezione d'urto totale per interazioni π^+p e la sezione d'urto totale per interazioni π^-p. I massimi e i minimi della curva, particolarmente evidenti a basse energie, indicano la presenza di risonanze. La curva a colori è un'estrapolazione a energie inferiori del comportamento regolare alle alte energie della curva in grigio. Il fatto che la curva a colori tenda a mediare i massimi e i minimi della curva in grigio viene chiamato dualità fenomenologica.



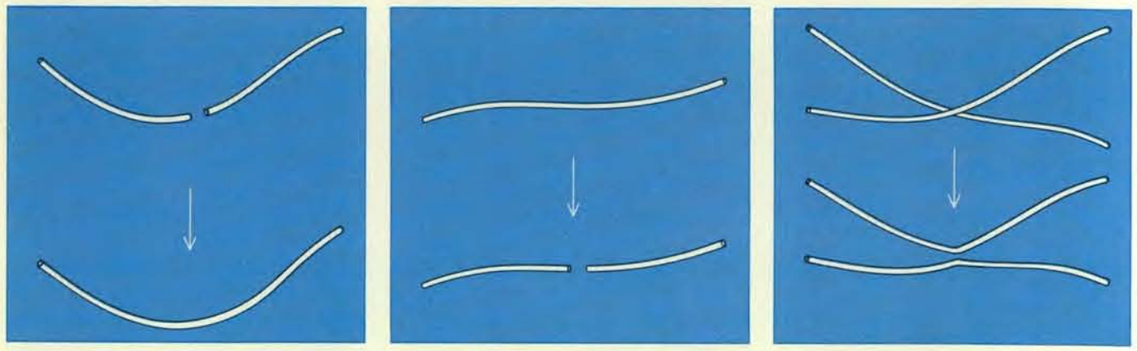
Due modi alternativi di interpretare le forze in gioco in un'interazione in cui due particelle elementari incidenti (*indicate con « a » e « b »*) si urtano per produrre due particelle uscenti (*indicate con « c » e « d »*) sono indicati in questo schema. Le forze si possono interpretare come il risultato di uno scambio di risonanze, come indicato nel diagramma schematico a sinistra, o come il risultato della formazione di una risonanza intermedia che successivamente decade, come nel processo di diffusione « di canale diretto » rappresentato a destra. La dualità teorica consiste nell'affermare che questi due processi di diffusione, trattati matematicamente in prima approssimazione, coincidono.



L'analogia delle corde senza massa si è rivelata utile nella descrizione degli stati quantici degli adroni nel modello a risonanza duale. Gli estremi della corda devono sempre muoversi con la velocità della luce. Sebbene una siffatta « corda di luce » sia supposta monodimensionale, la teoria dei quanti richiede che lo spazio nel quale la corda si muove nel più semplice dei modelli di risonanza duale debba avere 25 dimensioni! Per una data massa adronica la corda ha un numero finito di possibili moti nel sistema di riferimento del centro di massa, corrispondenti a un insieme di stati quantici discreti. Per esempio, la corda può ruotare attorno al proprio centro mantenendosi tesa (*in alto*), o contemporaneamente ruotare e oscillare (*in basso*).

dente conseguenza è che diventa impossibile soddisfare contemporaneamente i principi della meccanica quantistica e della relatività ristretta, a meno che le dimensioni dello spazio in cui si muovono le corde non abbiano un valore particolare. Nel mondo reale lo spazio ha ovviamente tre dimensioni: nel più semplice dei modelli a corde ne deve avere 25! Per fortuna i differenti modelli a corde hanno differenti dimensioni critiche, in modo tale che questa restrizione non rappresenta una difficoltà insormontabile. Piuttosto si intravvede l'interessante possibilità che il numero delle dimensioni dello spazio possa un giorno essere ricavato da altre grandezze, o almeno correlato ad esse.

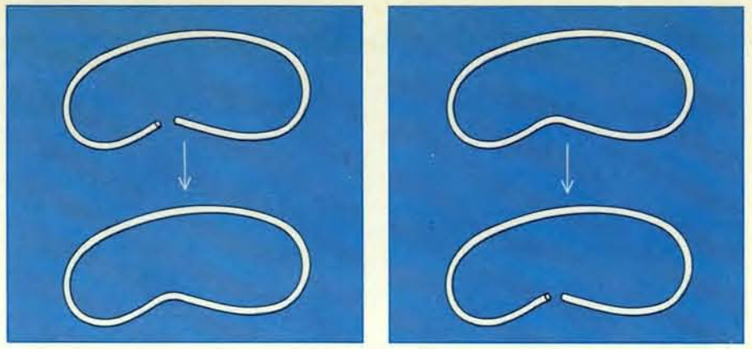
La relazione tra le particelle elementari e gli stati di moto di una ipotetica corda merita qualche considerazione ulteriore. Secondo le leggi della fisica classica una corda di luce ha un'infinità continua di moti possibili. La teoria dei quanti riduce il numero di possibilità nel sistema di riferimento del centro di massa a un insieme discreto di stati quantici. Anche il numero di stati quantici è infinito, ma fino a ogni dato valore della massa è finito: in effetti la distribuzione è esattamente del tipo di Hagedorn. Come possibile esempio di moto, supponiamo che la corda stia ruotando attorno al suo punto medio mantenendosi perfettamente tesa; in questo caso la frequenza di rotazione deve essere tale che il momento angolare sia un multiplo intero della costante di Planck ridotta. La lunghezza totale della corda è unicamente determinata in ciascun caso dalla condizione che i suoi estremi si muovano con la velocità della luce. Questa classe di stati contiene il massimo momento angolare per una data quantità di massa, perciò gli stati adronici descritti in questo modo corrispondono a quelli previsti dalla traiettoria di Regge. Vi sono molti altri possibili moti nei quali la corda subisce sia oscillazioni che rotazioni (*si vedano le figure nella pagina a fronte in alto*).



Le corde di luce possono interagire secondo la teoria della risonanza duale o congiungendosi (*a sinistra*), o rompendosi (*al centro*), o rompendosi e poi ricongiungendosi (*a destra*). Quest'ultimo caso è l'analogo del *crossing over* osservato nei cromosomi.

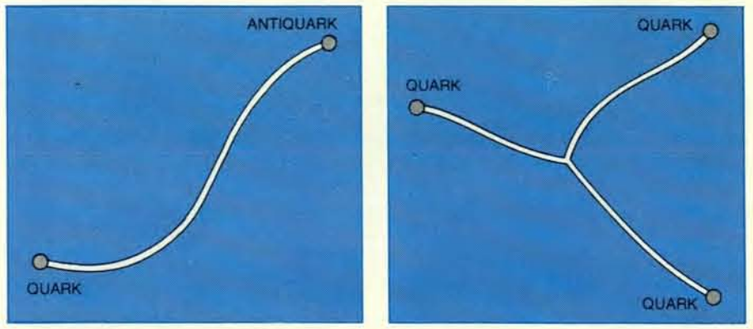
La teoria delle corde di luce esposta sopra descrive lo spettro di particelle di un particolare modello duale nella prima approssimazione. La situazione completa è più complessa, perché le corde possono interagire una con l'altra (*si vedano le figure nella pagina a fronte in basso*). Stanley Mandelstam dell'Università della California a Berkeley mostrò che la principale interazione può essere o la rottura di una corda in due pezzi o la giunzione di due corde per formarne una sola. Un'altra possibilità è che quando due corde si toccano in punti interni, esse si spezzino per poi ricongiungersi in modo che ogni segmento di una corda si congiunga a un segmento dell'altra (un fenomeno analogo in biologia è il doppio scambio di frammenti dei cromosomi.) Queste regole sull'interazione delle corde permettono di calcolare le ampiezze di diffusione per varie possibili reazioni degli adroni. Un effetto delle interazioni è di modificare in qualche modo le proprietà degli adroni stessi. Per esempio, le traiettorie di Regge non si mantengono esattamente rettilinee quando nell'esame delle interazioni si introducono correzioni al di là della prima approssimazione.

Quando si prendono in considerazione gli effetti dell'interazione delle corde, si scopre subito che è necessario un nuovo tipo di corda. Dato che due corde possono interagire congiungendosi agli estremi, che cosa può impedire ai due estremi di una corda di congiungersi, trasformandola in tal modo in una corda chiusa, o cappio? In effetti la consistenza matematica richiede che gli adroni a cappio siano inclusi nella teoria, anche se essi non sono previsti nella prima approssimazione (*si veda la figura sopra*). Anche gli adroni che corrispondono a stati quantici di corde chiuse giacciono su traiettorie di Regge parallele rettilinee (nella prima approssimazione), con una pendenza che è esattamente la metà di quella delle traiettorie associate alle corde aperte. La possibile esistenza di questi nuovi stati nella teoria è eccitante, poiché i fatti sperimentali richiedono qualcosa in più delle ordinarie traiettorie di Regge a corde aperte. Il comportamento alle alte energie delle reazioni di diffusione elastica ha mostrato un tipo di interazione diffrattiva. Un punto di vista ampiamente sostenuto è che la diffusione di diffrazione può essere spiegata in termini di scambi singoli e multipli della traiettoria di Regge di un « pomeron » (il nome « pomeron » deriva da quello del fisico russo Isaak Yakovlevich Pomeranchuk). Le corde chiuse costituiscono dei validi candidati per le traiettorie pomeroniche. Infatti, questa particolare teoria richiede che le particelle siano associate a traiettorie pomeroniche. Finora non è stato rivelato in laboratorio nessun serio candidato al ruolo di particella pomeron, ma vi sono parecchi motivi per ritenere che tali particelle siano difficilmente rivelabili.

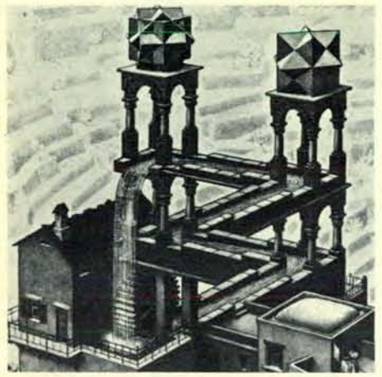


Un nuovo tipo di corda di luce necessario per la consistenza matematica del modello, è una corda chiusa, o cappio, che si forma allorché una corda aperta si congiunge agli estremi (*a sinistra*). Viceversa, dovrebbe essere possibile anche la rottura di una corda chiusa per formare una corda aperta (*a destra*). Particelle ipotetiche chiamate pomeroni sembrano essere descrivibili mediante il modello di risonanza duale a corda chiusa, ma tali particelle non sono ancora state osservate sperimentalmente in laboratorio.

Dove si inseriscono i quark nel modello a corde? Un'interessante possibilità è che essi siano particelle puntiformi (come in qualsiasi teoria dei campi) fissate agli estremi delle corde. Se si suppone che compaiano nuovi quark ogniqualvolta si formano nuovi estremi per la rottura delle corde, allora questo quadro fornisce una semplice spiegazione del confinamento dei quark:



I quark possono essere inseriti in molti modi possibili nei modelli di risonanza duale. Per esempio, se si suppone che i quark siano particelle fondamentali puntiformi fissate agli estremi delle corde di luce, sarebbe facile descrivere (*a sinistra*) un adrone costituito da un quark e da un antiquark (cioè un mesone), mentre un adrone costituito da tre quark (cioè un barione) richiederebbe una topologia più complessa (*a destra*).



Primo Levi
Il sistema periodico

Gli incontri, gli imprevisti, le sfide del mestiere di chimico, raccontati attraverso ventun «momenti», ognuno dei quali prende spunto da un elemento: l'azoto, il carbonio, il nichel, e cosí via. Un «romanzo della materia» che è anche la storia di una formazione civile, negli anni a cavallo della guerra. Lire 2600.

Einaudi

i quark sono tenuti assieme da corde! A prima vista il quadro appare promettente per la classe di adroni chiamati mesoni, che contengono un quark e un antiquark e per gli stati pomeronici, che non contengono quark. Gli adroni chiamati barioni, invece, sono composti da tre quark ciascuno e richiedono quindi una costruzione più complicata, quale una corda a forma di Y, in modo da presentare tre estremi liberi ai quali fissare i quark (*si veda la figura nella pagina precedente in basso*). Questa topologia dà origine a serie difficoltà matematiche; in effetti anche la costruzione dei mesoni presenta certi problemi. I quark dovrebbero avere una carica elettrica e uno spin, ma pare che tali particelle non possano attaccarsi saldamente agli estremi di una corda in moto con la velocità della luce. Si potrebbe anche cercare di rallentare gli estremi attribuendo ai quark una massa, ma questo solleva altri seri problemi. Nonostante l'interesse iniziale, attaccare i quark agli estremi di corde può rivelarsi un metodo completamente errato per incorporarli nei modelli a risonanza duale.

Un'alternativa all'idea di fissare i quark agli estremi di corde potrebbe essere che i vari numeri quantici (come la carica elettrica) che si suppone che i quark trasportino, siano distribuiti sulla lunghezza delle corde. In questo caso i barioni non richiedono necessariamente una nuova topologia. Questa ipotesi che i numeri quantici dei quark non siano localizzati in punti ben precisi è alquanto impopolare fra i fisici teorici. Il principale motivo sta nel fatto che recenti esperimenti eseguiti al Centro dell'Acceleratore Lineare di Stanford (SLAC) sulla diffusione ad alta energia di elettroni e protoni mostrano la proprietà chiamata *scaling* (invarianza di scala).

Lo scaling è solitamente interpretato come una prova diretta del fatto che la carica elettrica all'interno di un protone è in un dato istante localizzata in uno o più punti, che corrispondono alle posizioni dei quark. Un'altra alternativa che dovrebbe essere compatibile con lo scaling è che i quark siano particelle puntiformi mobili lungo la corda come le perle di una collana. Sfortunatamente questo modo di affrontare il problema porta a una matematica molto complessa, e quindi le sue previsioni non possono in alcun modo essere verificate.

Una lacuna più grave del più semplice dei modelli a corde, oltre a quella di essere vincolato a 25 dimensioni, è la sua incapacità di giustificare l'esistenza dei fermioni, particelle con spin semiintero (lo spin, come le altre proprietà quantistiche, può avere solo certi valori discreti). Tra gli adroni queste particelle sono esattamente quelle chiamate barioni. Se viene rifiutata l'ipotesi di poter ricavare uno spin semiintero attaccando a una corda tre quark, ciascuno con mezza unità di spin, allora è la corda stessa che deve in qualche modo dare origine allo spin. Stati di fermioni duali non interagenti furono proposti teoricamente da Pierre Ramond nel 1971. Poco tempo dopo fu messo a punto da André Neveu e dal sottoscritto un modello duale di particelle interagenti dotate di spin che comprendeva i fermioni di Ramond. Tale modello soddisfa gli stessi principi generali del più semplice di tali modelli, ma prevede uno spettro di particelle più realistico. È stato dimostrato che una descrizione a corde di questo modello è in accordo con la relatività ristretta e con la meccanica quantistica, purché lo spazio abbia nove dimensioni. Anche se questo non è ancora il numero giusto, esso costituisce un passo in avanti nella direzione giusta. Pare inoltre che esista almeno la possibilità che sei delle nove dimensioni possano essere attribuite a gradi di libertà interni anziché allo spazio ordinario.

Se si riuscisse a fare ciò in un modo pienamente autoconsistente, allora forse una forma modificata del modello potrebbe essere quella giusta.

Il modello duale con spin rappresenta un'estensione non trascurabile del precedente modello duale. Mentre la «azione» che caratterizza il modello senza spin presenta un gruppo di simmetrie matematiche analogo a quello della teoria della relatività generale, il modello con spin presenta un gruppo di simmetrie ancora più vasto. Se l'eleganza di una teoria viene misurata dal numero di simmetrie che essa possiede, allora il modello duale con spin ha un'alta quotazione. La mia opinione è che l'eleganza, definita in questo modo, sia strettamente correlata all'importanza fisica.

Per concludere, negli ultimi sei anni è stata sviluppata una nuova teoria delle particelle elementari – la teoria della risonanza duale. Essa non ha ancora fornito nuove previsioni di fatti sperimentali di rilievo che possano confermarne la validità (o escluderla), ma essa rende conto in modo soddisfacente di molti aspetti generali del sistema degli adroni.

Se si potranno superare i pochi ostacoli che ancora rimangono, vi sono buone probabilità che questa ricerca possa alla fine tradursi in una teoria completa delle interazioni forti e delle particelle che a esse sono soggette.

Il mantello terrestre

Questo grande involucro di roccia incandescente comprende l'80 per cento del volume della Terra e il 67 per cento della sua massa. Per quanto il mantello sia inaccessibile, si è appreso molto su di esso attraverso esperienze indirette

di Peter J. Wyllie

Il mantello terrestre è uno spesso guscio di roccia incandenscente che separa il nucleo della Terra, metallico e parzialmente fuso, dalle rocce più fredde della sottile crosta. Esso ha inizio a una profondità che va in media da 35 a 45 chilometri sotto la superficie e prosegue fino ad una profondità di circa 2900 chilometri; interessa così circa la metà del raggio della Terra, l'83 per cento del suo volume e il 67 per cento della sua massa. Ha una profonda influenza sulla crosta: questa infatti, con la sua sottile pellicola di oceani e di atmosfera, è un distillato del mantello dal quale hanno origine le forze propulsive che muovono lentamente i continenti attorno alla superficie terrestre. La conoscenza del mantello è quindi fondamentale per la comprensione della struttura e del comportamento dinamico della Terra. Anche se il mantello è inaccessibile, è stata raccolta su di esso, con mezzi più o meno indiretti, una considerevole messe di informazioni.

Il mantello esercita la sua influenza sulla superficie in molti modi. Per esempio, nei 4,6 miliardi di anni trascorsi da quando la Terra ha avuto origine dalla nebulosa solare, il mantello, per fusione dei suoi costituenti a più basso punto di fusione, ha prodotto lave che hanno raggiunto la superficie e si sono solidificate, aggiungendo altre rocce alla crosta e liberando vapore acqueo e altri gas nell'atmosfera e negli oceani. Su scala diversa i composti gassosi del carbonio provenienti dal mantello hanno dato inizio alla storia della vita, fornendo la materia prima per le molecole organiche.

Analogamente, il mantello è sede di una forza propulsiva che ha molteplici effetti: muovendosi lentamente sotto la crosta, il mantello plasma la superficie terrestre; determina la nascita e la persistenza delle montagne che, in sua assenza, verrebbero spianate dall'erosione fino al livello del mare in meno di 100 milioni di anni; provoca le eruzioni vulcaniche, i terremoti e la deriva dei continenti.

Una sezione trasversale della Terra ne mostra gli strati concentrici del nucleo, del mantello e della crosta (*si veda la figura in alto nella pagina a fronte*). Essi differiscono l'uno dall'altro per la composizione o per lo stato fisico, o anche per entrambi. Il mantello è composto di minerali silicatici ricchi in magnesio e ferro, e ha una composizione media che corrisponde a quella di una roccia chiamata peridotite dal nome del suo componente principale, l'olivina, usata come gemma e detta anche peridoto nella sua forma verde e trasparente. Il mantello è solido, ma in una zona relativamente sottile che va da circa 100 chilometri sotto la superficie a circa 250 chilometri, la roccia può essere parzialmente fusa, con sottili pellicole di liquido distribuite fra i granuli minerali. Questa zona, per ragioni che verranno chiarite più avanti, viene detta strato a bassa velocità.

La densità del mantello aumenta con la profondità da circa 3,5 grammi per centimetro cubo presso la superficie a circa 5,5 grammi presso il nucleo. L'aumento non è continuo; la curva della densità mostra dei gradini ben distinti (*si veda la figura a pagina 42*). Essi indicano cambiamenti significativi nelle rocce del mantello a profondità vicine a 400 e 650 chilometri. La distribuzione delle densità offre la base per calcolare che il mantello costituisce il 67 per cento della massa della Terra.

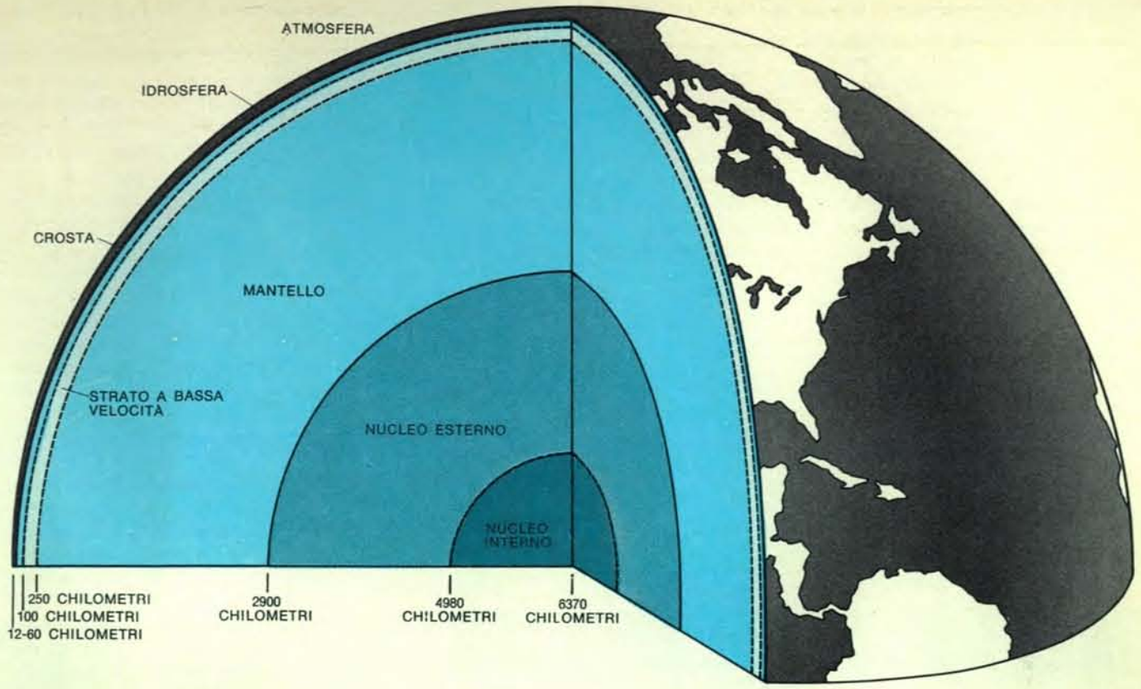
Tettonica a zolle

Questo quadro statico di una Terra concentrica e stratificata viene modificato dalla teoria della tettonica a zolle, che si riferisce al moto delle zolle di litosfera. La litosfera comprende la crosta e parte del mantello superiore, e viene distinta dall'astenosfera, che le sottostà, in base al fatto che è più fredda, e quindi più rigida. La teoria della tettonica a zolle offre un quadro dinamico di un mantello mobile, con zolle di litosfera spesse circa 100 chilometri che si spostano sopra l'astenosfera. La superficie della Terra è coperta da poche grandi zolle litosferiche e parecchie altre minori. Queste zolle, simili a gusci, si muovono le une rispetto alle altre e l'attività geologica, testimoniata da terremoti e vulcani, si concentra lungo i limiti che dividono una zolla dall'altra, cioè lungo i margini di zolla.

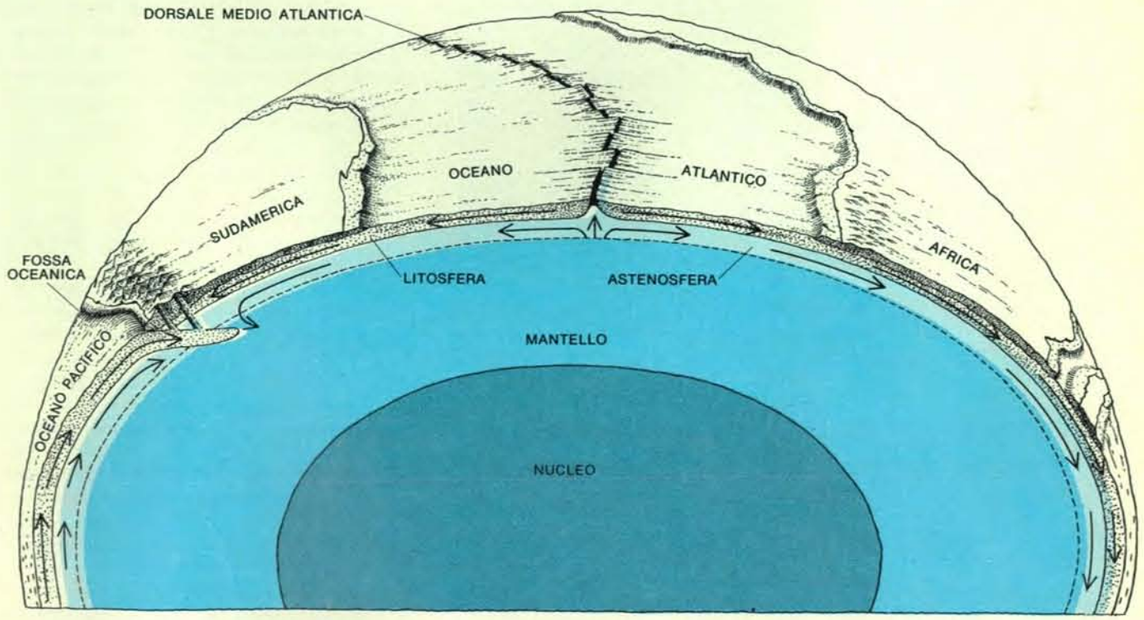
I margini delle zolle sono di tipo divergente o convergente. Sotto la cresta delle dorsali oceaniche risale il materiale dall'astenosfera, fondendo a mano a mano che si muove e producendo così della lava che viene eruttata nella valle tettonica che si estende al centro della dorsale; in questo modo si genera nuova crosta. I moti convettivi presenti nel mantello fanno divergere le zolle là dove si forma nuova crosta.

Dove i loro margini convergono le zolle possono collidere, sollevando la crosta fino a formare montagne, oppure una zolla può spingersi sotto l'altra, riportando nel mantello il materiale della litosfera; di conseguenza la nuova litosfera che si forma è compensata da quella che si distrugge altrove. Tali margini sono associati alle fosse oceaniche e agli allineamenti di vulcani, compresi gli archi di isole vulcaniche e i vulcani nelle catene montuose attive.

Della Terra nel suo insieme sono state determinate molte proprietà fisiche: le dimensioni, la forma e la massa del nostro pianeta sono state misu-



L'interno della Terra delineato tenendo conto solo degli strati e non dei processi attivi che hanno luogo nel suo interno. Le rocce della sottile crosta sono fredde e rigide. La roccia del mantello, che è caldo, è in grado di muoversi lentamente. I dati delle onde sismiche indicano che il nucleo esterno consiste di metallo fuso. L'idrosfera è costituita dalle acque di superficie.



Sezione trasversale della Terra dinamica, come viene vista dalla teoria della tettonica a zolle. Le zolle di litosfera, la quale comprende la crosta e parte del mantello superiore, si spostano lateralmente sull'astenosfera, che è uno strato del mantello caldo e forse parzialmente fuso. Il materiale dell'astenosfera risale sotto le creste delle dorsali oceaniche, fonde e produce della lava che viene eruttata e va a formare nuova crosta nel fondo oceanico. Le zolle di litosfera divergono a mano a mano che il materiale che risale forma della nuova litosfera; la formazione di nuova litosfera viene bilanciata dalla distruzione di una quantità equivalente di litosfera ai margini delle zolle che convergono, dove la piastra di litosfera si spinge giù nel mantello. Questo tipo di margine è associato in modo caratteristico sia alle fosse oceaniche sia agli allineamenti di vulcani.

rate con grande precisione. In base al suo volume e alla sua massa si può calcolare una densità media di 5,5 grammi per centimetro cubo, molto più alta, cioè, di quella delle rocce che costituiscono la parte accessibile della crosta. Una parte notevole del suo interno, quindi, dev'essere composta di un materiale a densità più alta di 5,5 grammi per centimetro cubo. Gli studi sul campo gravitazionale terrestre e sulle proprietà delle sfere rotanti indicano che la massa è concentrata verso il centro.

Dati sul mantello

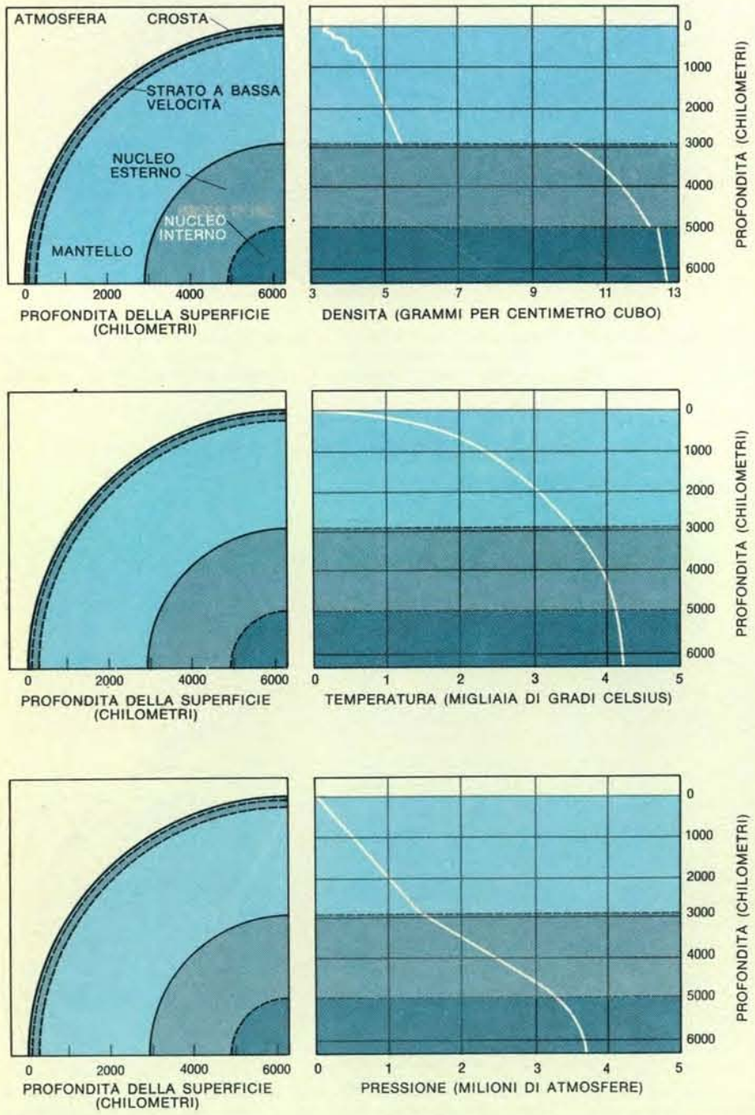
La principale fonte di informazioni sulla fisica della Terra proviene dagli studi sulle onde sismiche, le quali forniscono qualcosa di simile a una radiografia dell'interno della Terra. Quest'ultima, infatti, quando viene scossa da un terremoto violento, risuona come una campana; le vibrazioni di una campana dipendono dalla sua forma e dalle sue proprietà fisiche, e, analogamente, le vibrazioni della Terra, registrate da strumenti sensibili, possono venire interpretate in modo da avere indicazioni sulle sue proprietà.

La liberazione di energia al fuoco (ipocentro) di un terremoto genera diversi tipi di onde: sia le onde primarie *P* sia le onde secondarie *S*, più lente, si propagano all'interno della Terra. Le abbreviazioni servono anche a ricordare che l'energia viene trasmessa in due modi diversi lungo un certo percorso: le onde *P* sono onde di compressione e rarefazione, mentre le onde *S* sono dovute a deformazioni di taglio (*si veda la figura nella pagina a fronte*). Le onde *P* attraversano sia i solidi sia i liquidi, le *S*, invece, vengono trasmesse soltanto dai materiali che possono subire deformazioni di taglio, cioè materiali che possono venire piegati o deformati. Un'onda *S* non può attraversare un liquido perché il liquido non offre resistenza a deformazioni di taglio e scorre, invece, molto facilmente.

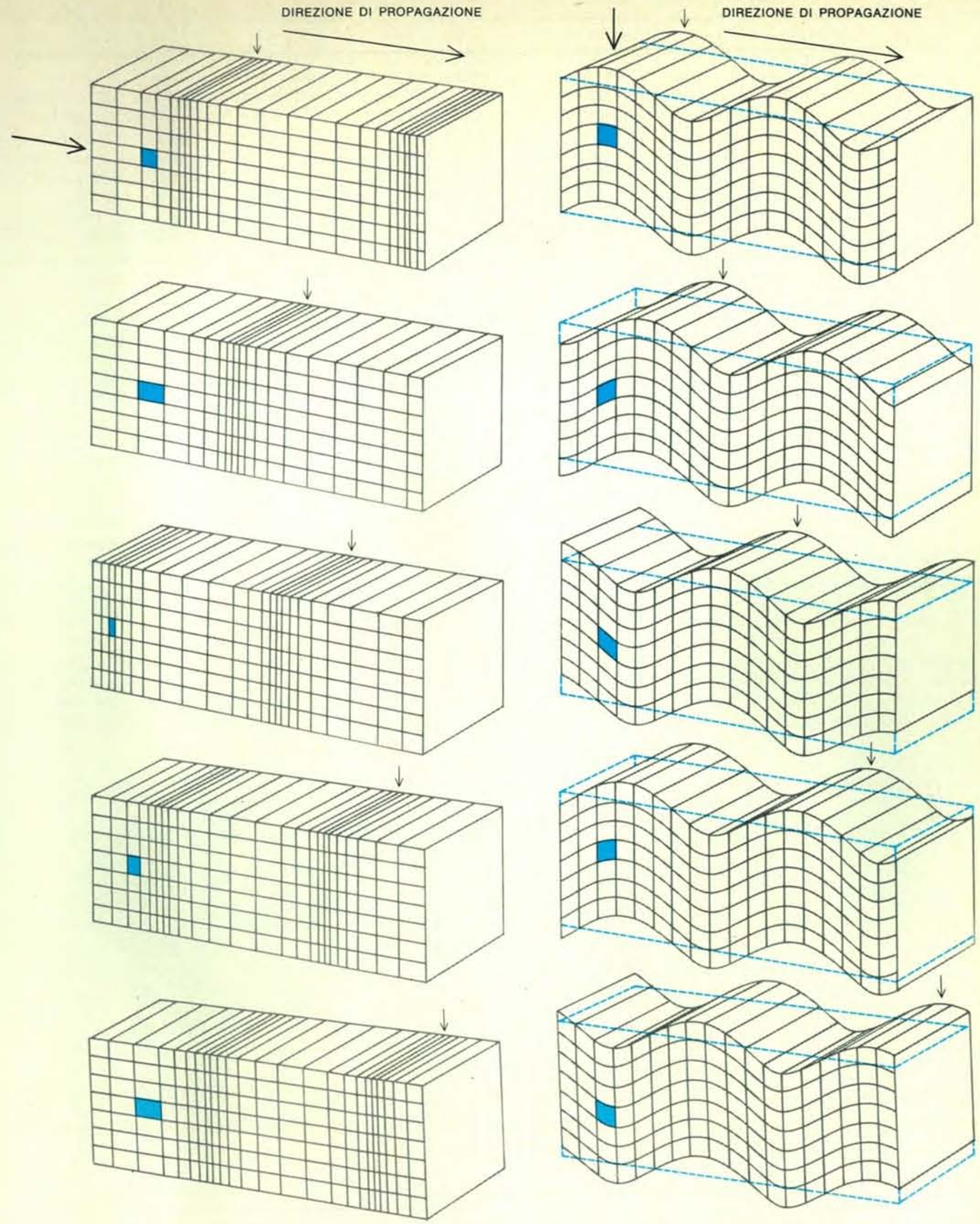
Se le proprietà del materiale che compone la Terra fossero uniformi dappertutto, le onde, uscendo dal fuoco di un terremoto, seguirebbero linee rette e ogni tipo di onda si propagherebbe a velocità costante. Il tempo impiegato dalle onde *P* e *S* a raggiungere una certa stazione di registrazione, posta a distanza nota dal fuoco, darebbe la velocità di ogni tipo di onde. In realtà, invece, i risultati ottenuti da molti terremoti indicano che le onde viaggiano all'interno della Terra a velocità superiore a quella che si potrebbe prevedere in base alle velocità che si conoscono nelle rocce superficiali. Inoltre questi risultati indicano che le onde che percorrono distanze maggiori viaggiano anche a velocità superiore. Questi dati significano che la velocità delle onde sismiche a grande profondità è maggiore che presso la superficie e che essa aumenta progressivamente con la profondità.

In base a queste e ad altre osservazioni è noto che le onde sismiche all'interno della Terra vengono riflesse e rifratte. In seguito alla rifrazione esse seguono dei percorsi concavi verso l'alto. Si ha riflessione a quelle profondità alle quali si hanno dei cambiamenti improvvisi nelle proprietà fisiche. La riflessione si verifica quindi ai limiti: fra crosta e mantello, fra mantello e nucleo, fra nucleo esterno e nucleo interno. Queste caratteristiche delle onde sismiche testimoniano la struttura concentrica della Terra; le onde *S* seguono percorsi che non vanno oltre il limite mantello-nucleo: non attraversano cioè il nucleo, e questo dato dimostra che almeno la parte esterna del nucleo è liquida.

La misura dei tempi di viaggio delle onde lungo diversi percorsi nel man-



Le proprietà fisiche della Terra variano con la profondità della superficie. Le profondità sui grafici della densità, della temperatura e della pressione corrispondono alle profondità sulle sezioni disegnate a sinistra. Un'atmosfera corrisponde alla pressione che l'aria esercita al livello del mare e vale 1033 grammi per ogni centimetro quadrato.



Onde sismiche di due tipi si propagano nell'interno della Terra a partire dal fuoco di un terremoto, dove si libera dell'energia. Esse sono qui rappresentate mentre attraversano un blocco di roccia. A sinistra è un'onda di compressione, o *P*, che ha origine da una pressione o una rarefazione improvvisa in direzione della propagazione dell'onda. Questa comprime la roccia e le particelle vicine si muovono in avanti; poi rimbalzano indietro alla posizione iniziale e oltre, continuando a vibrare in questo modo per qualche tempo. Ogni volumetto di materia (*colore*) si contrae e si espande man mano che l'onda di compressione si muove nella roccia. Un'onda trasversale, o *S*, ha origine da una pressione ad angolo retto con la direzione di propagazione. Le particelle di roccia vibrano in su e in giù; un pezzetto di materiale (*colore*) subisce deformazioni di taglio. Il fronte d'onda trasmette energia. Questa illustrazione è stata adattata da *La Geofisica*, di O. M. Phillips pubblicata in Italia da Mondadori Est.

tello fornisce un modo per calcolare a che velocità si trasmettono le onde nel materiale che compone il mantello alle varie profondità. I diagrammi della velocità delle onde sismiche mostrano un aumento progressivo con la profondità, ma fino a circa 1000 chilometri gli aumenti avvengono a salti. Queste osservazioni indicano che il mantello ha una struttura stratificata. La velocità sia delle onde *P* sia delle onde *S* nel mantello diminuisce in una zona che va circa da 100 a 250 chilometri dalla superficie. Questo è lo strato a bassa velocità a cui si è già accennato: nel modello della tettonica a zolle è considerato equivalente all'astenosfera. Le proprietà fisiche di questo strato, dedotte dai grafici della velocità delle onde e da altri dati geofisici, sono in accordo con la presenza di un liquido interstiziale, che in genere è ritenuto una frazione fusa della roccia stessa che compone lo strato.

Margini di zolla

I terremoti ci offrono non solo questo quadro statico di una Terra a strati concentrici, ma anche alcuni aspetti del quadro dinamico della tettonica a zolle. Poiché i terremoti avvengono soltanto in rocce che siano abbastanza fredde e rigide da potersi fratturare, la distribuzione dei loro fuochi indica i margini delle zolle stabili. La profondità dei fuochi che si trovano lungo i margini associati a dorsali oceaniche è inferiore ai 100 chilometri, e questo indica che il mantello al di sotto della litosfera è caldo.

Presso i margini convergenti, dove una zolla sprofonda nel mantello al di sotto di un'altra, i fuochi dei terremoti si spingono fino a profondità attorno ai 700 chilometri. La distribuzione di questi fuochi profondi permette di individuare i frammenti di litosfera che affondano; studi dettagliati sulla velocità delle onde sismiche in queste zone confermano l'esistenza di piastre di litosfera più fredda del mantello che si estendono dentro di esso fino a profondità considerevoli.

Le informazioni sulla struttura e sulle proprietà fisiche del mantello che ci vengono date dalle onde sismiche sono abbastanza dirette; al contrario, per ottenere delle informazioni sulle proprietà chimiche dell'interno della Terra, bisogna ricorrere essenzialmente a dati indiretti. I soli campioni chimici disponibili sono alcune rocce giunte nella crosta dagli strati più superficiali del mantello. Per poter valutare la composizione della Terra nel suo insieme, e quella del nucleo e del mantello, bisogna affidarsi alla chimica dei corpi extraterrestri, ivi compresi le meteoriti, il sole e le stelle. Si suppone che la composizione di questi corpi sia sostanzialmente simile a quella della Terra. Questo modo di affrontare il problema comporta la formulazione di modelli fisici e chimici sull'origine del sistema solare e della Terra.

Si accetta, in genere, che il sistema solare si è formato 4,6 miliardi di anni fa per il collasso gravitazionale di materia che prima si trovava dispersa nello spazio interstellare. Un grande precursore del Sole, un proto Sole, era circondato da una sottile nebulosa, a forma di disco, formata di gas e particelle di polvere. L'aggregazione locale di particelle e la condensazione di gas all'interno della nebulosa rotante formarono dei piccoli oggetti che, aggregandosi, diedero luogo a corpi planetari. In questo periodo si formò il materiale delle meteoriti. Poiché il Sole comprende più del 96,6 per cento della massa del sistema solare, la sua composizione è realmente la stessa di tutto il sistema solare. La concentrazione degli elementi nel Sole e nelle altre stelle è stata misurata con metodi spettroscopici, in base al fatto che ogni elemento viene identificato dalla sua radiazione elettromagnetica caratteristica.

Cosa rivelano le meteoriti

Le meteoriti viaggiano adesso attraverso il sistema solare su orbite ellittiche che di tanto in tanto intercettano la Terra. Esistono prove che esse provengono dalla fascia degli asteroidi, quello sciame di piccoli corpi planetesimali che orbitano a distanze fra 2,2 e 3,2 unità astronomiche dal Sole (si veda anche l'articolo *La natura degli asteroidi* su « Le Scienze », n. 81, maggio 1975). Le meteoriti sono molto variabili per composizione chimica, per mineralogia e per struttura, ma per lo scopo di questa discussione sarà sufficiente distinguerne i due gruppi principali: le meteoriti ferrose e le meteoriti litoidi. Le meteoriti ferrose sono costituite essenzialmente da leghe di ferro-nichel, in cui la concentrazione di nichel varia tra il 4 e il 20 per cento; contengono anche piccole quantità di solfuro di ferro.

Le meteoriti litoidi sono composte essenzialmente da minerali silicatici, assieme a leghe metalliche e solfuro di ferro in varie proporzioni. Le abbondanze relative di elementi non volatili quali il magnesio, il calcio e il ferro sono più o meno le stesse, in molti tipi di meteoriti litoidi, che si hanno nel Sole e nelle altre stelle. Se ne deduce quindi che queste abbondanze forniscono una buona base per la stima delle abbondanze complessive nella Terra e negli altri pianeti.

Nei modelli complessi sull'origine e sull'evoluzione del sistema solare la composizione della Terra viene ricavata a partire da una meteorite litoide ricca in elementi volatili e immaginando una serie di processi e di cambiamenti chimici che potrebbero spiegare gli altri tipi di meteoriti e l'attuale struttura della Terra. Questi calcoli permettono delle stime della composizione del nucleo e del mantello.

Un metodo d'indagine più semplice, che fornisce risultati simili, è di scegliere un certo gruppo di meteoriti litoidi e di supporre che la composizione media della loro frazione silicatica sia equivalente a quella del mantello terrestre. I solfuri di ferro e una parte sufficiente delle leghe ferro-nichel delle stesse meteoriti forniscono una stima della composizione del nucleo terrestre. Le meteoriti ferrose servono come controllo per questi calcoli.

Qualsiasi procedimento venga seguito, le stime della composizione del mantello sono d'accordo sui seguenti punti: 1) più del 90 per cento (in peso) del mantello consiste di ossidi di silicio, magnesio e ferro (SiO_2, MgO e FeO), e nessun altro ossido supera il 4 per cento; 2) gli ossidi di alluminio (Al_2O_3), calcio (CaO) e sodio (Na_2O) in totale oscillano fra il 5 e l'8 per cento; 3) questi sei ossidi rappresentano più del 98 per cento del mantello, e nessun altro ossido raggiunge una concentrazione dello 0,6 per cento. Le concentrazioni degli altri elementi, presenti in tracce, non vengono definite. Nelle rocce del mantello gli ossidi sono combinati in minerali diversi.

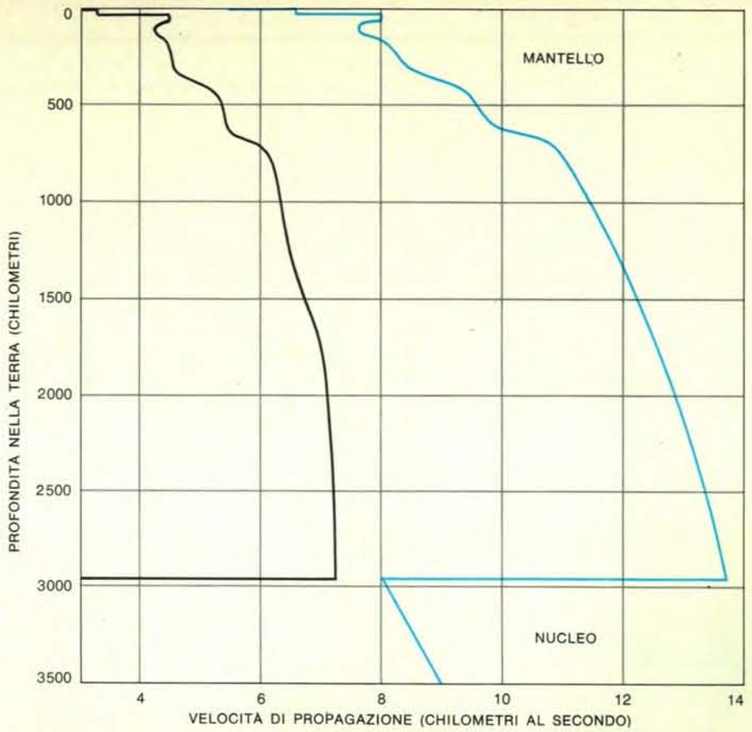
Solo le peridotiti, fra tutte le rocce che si trovano nella crosta, corrispondono alle stime della composizione del mantello eseguite attraverso lo studio dei corpi extraterrestri. È dunque a tali peridotiti che si rivolge l'attenzione dei geologi che si interessano al mantello e che desiderano altri dettagli, fra cui le concentrazioni degli elementi in traccia. Il problema, ovviamente, è di assicurarsi che il campione di peridotite che si sta studiando si sia formato nel mantello e non nella crosta. Per questa ragione hanno un interesse particolare quegli inclusi arrotondati, detti noduli, che si trovano nei camini (pipes) kimberlitici (intrusioni cilindriche, del diametro di qualche centinaio di metri, che in certe zone perforano la crosta e che provengono dal mantello).

Kimberliti e lave

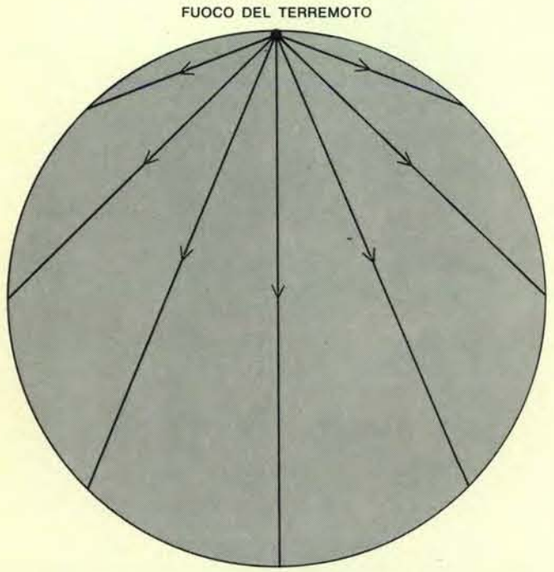
Le kimberliti sono celebri come rocce che portano i diamanti alla superficie terrestre. Il diamante è una fase del carbonio stabile soltanto a pressioni elevatissime; quindi si può essere certi che le kimberliti hanno origine nell'intervallo compreso fra i 150 e i 300 chilometri di profondità dalla superficie – ben dentro il mantello superiore.

Diversi fatti indicano che un camino kimberlitico, al momento della sua formazione, risale rapidamente attraverso la crosta come sospensione fluida di solidi, roccia fusa e gas, irrompendo alla superficie con un tremendo scoppio e dando luogo a una breve esplosione vulcanica. Fra i pezzi di roccia strappati dal condotto del camino e riportati alla superficie si hanno noduli di peridotite e, meno frequentemente, noduli di eclogite, un'altra roccia del mantello. Questi noduli sono arrotondati e lucidati per gli urti continui con gli altri frammenti spinti dal gas nel camino esplosivo.

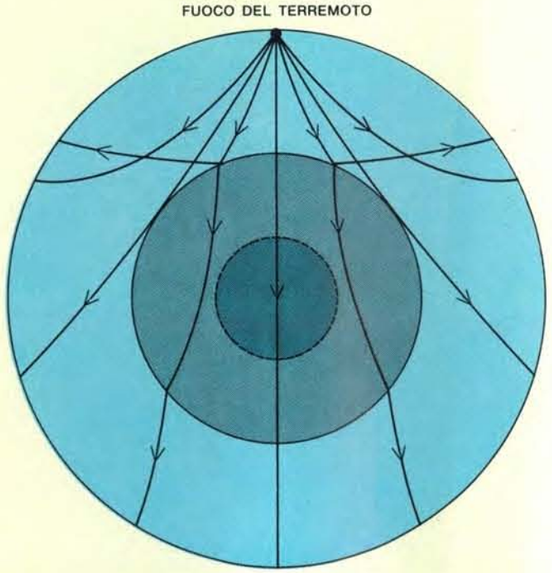
Anche in certe lave vulcaniche si hanno delle associazioni di noduli simili a quelle che si trovano nelle kimberliti. In genere questi noduli provengono da livelli del mantello più superficiali di quelli delle kimberliti. Molte



Diagrammi di velocità delle onde sismiche che attraversano il mantello, disegnati per le onde *S* (*in nero*) e *P* (*in colore*). Le velocità risentono dell'aumento di temperatura e pressione con la profondità, ma i gradini che appaiono a intervalli fino a profondità di circa 1000 chilometri corrispondono a cambiamenti nelle proprietà fisiche del mantello. La diminuzione di velocità fra 100 e 250 chilometri probabilmente è dovuta alla presenza di roccia parzialmente fusa: questa è l'astenosfera mobile della tettonica a zolle.



I percorsi delle onde sismiche nell'interno della Terra danno delle informazioni sulla sua struttura e sulle proprietà fisiche dei suoi strati concentrici. Dal fuoco di un terremoto l'energia si propaga in tutte le direzioni. Se le proprietà della Terra fossero dappertutto uniformi i percorsi delle onde seguirebbero linee rette (*a sinistra*) e le velocità delle onde sarebbero costanti.



Le misure dei tempi di viaggio mostrano invece che la velocità delle onde, e quindi le proprietà fisiche, cambiano improvvisamente a certi livelli (*a destra*), rivelando così gli strati concentrici. Le onde *P* vengono trasmesse sia dai solidi che dai liquidi. Le onde *S*, non trasmesse dai liquidi, non attraversano il nucleo: almeno il nucleo esterno della Terra è allo stato liquido.

stime della composizione del mantello si sono basate sulla mineralogia e sulla chimica delle associazioni di noduli delle kimberliti e delle lave. A questo scopo sono state usate anche peridotiti che si trovano in altri ambienti geologici, per quanto con queste ultime sia particolarmente importante tenere conto della natura dell'associazione geologica con altre rocce e dei risultati dei processi geologici, per essere ben certi che i risultati ottenuti in base alle misure eseguite sulla roccia si riferiscano davvero al mantello e non a fenomeni della crosta.

Oltre allo studio dei corpi extraterrestri e delle rocce del mantello, un terzo mezzo di indagine è quello di calcolare la composizione chimica di una peridotite immaginaria adatta a fornire le lave vulcaniche che derivano dal mantello per fusione parziale. La fusione dà luogo al magma basaltico che viene eruttato abbondantemente alla superficie dai vulcani e lascia nel mantello dietro di sé una peridotite residuale, presumibilmente impoverita degli elementi a più basso punto di fusione che, fondendo, sono stati asportati dal magma. Assegnando una composizione chimica adeguata alla peridotite residuale si arriva a una ipotetica composizione del mantello superiore; questa peridotite ipotetica è stata chiamata pirolite (roccia a pirosseno e olivina).

Il mantello eterogeneo

Un tempo si supponeva, sulla base di modelli fisici, che il mantello superiore avesse una composizione omogenea. Invece, come abbiamo visto, gli studi geofisici dettagliati degli ultimi anni indicano una struttura stratificata. In effetti i noduli che sono stati studiati confermano il fatto che il mantello superiore è molto eterogeneo, sia da un punto di vista chimico che mineralogico. I campioni a disposizione provengono da una zona del mantello che va dal limite con la crosta alla zona di origine dei camini kimberlitici, da un intervallo di profondità, cioè, fino a 250 chilometri. Si potrebbe sperare di trovare, fra i noduli, campioni della peridotite originale del mantello, che non sia mai stata fusa; campioni di peridotite residuale impoverita, dalla quale, cioè, è stata asportata una frazione fusa dei componenti a più basso punto di fusione; campioni di questa frazione fusa che non sono riusciti a raggiungere la superficie come magma, ma sono invece cristallizzati nel mantello sotto alta pressione formando della eclogite; campioni intermedi fra i tipi accennati e campioni dalla stessa composizione mineralogica, ma formati da altri processi, troppo complicati perché si possano prendere in considerazioni qui.

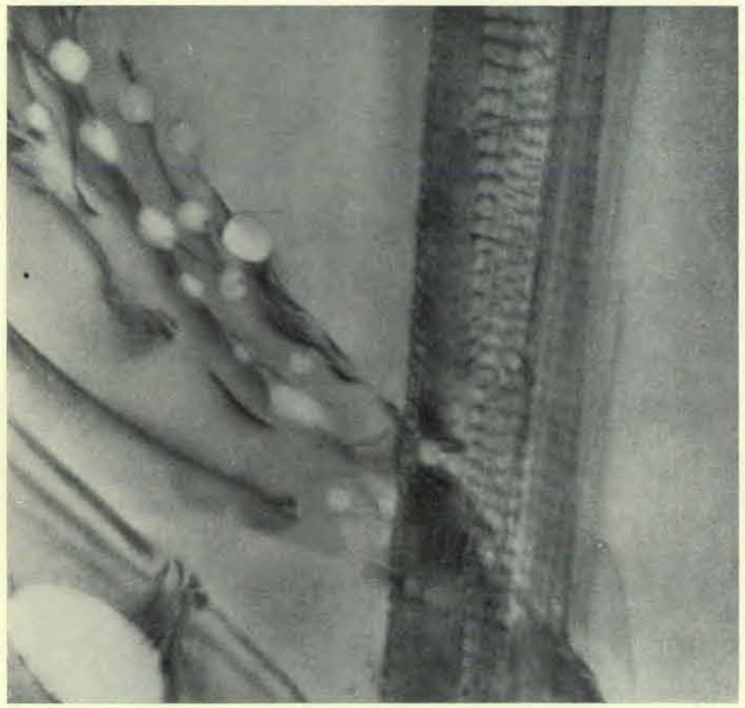
Le stime sulla composizione del mantello nel suo complesso, basate sull'esame dei corpi extraterrestri, forniscono dei valori simili a quelli che si ottengono dallo studio delle rocce che hanno origine nel mantello superiore. Queste stime, nel loro insieme, fanno pensare che il mantello, dall'alto in basso, come composizione, non cambi molto. Paragonando queste stime, però, si trova che la concentrazione di potassio del mantello superiore nel modello della pirolite è notevolmente più alta di quella che si trova nelle rocce peridotitiche che derivano dal mantello, e entrambe sono molto più basse delle stime ottenute dallo studio dei corpi extraterrestri.

Rimangono ancora grandi incertezze su concentrazione e distribuzione di altri elementi in traccia e di componenti volatili come l'acqua e l'anidride carbonica. Questi componenti hanno un significato fondamentale per definire la quantità di energia termica generata per decadimento radioattivo (di elementi come l'uranio, il torio e l'isotopo radioattivo del potassio, il potassio-40), la temperatura di fusione (in presenza di piccole quantità di acqua ad alta pressione le rocce cominciano a fondere a temperature più basse del punto di fusione della roccia asciutta) e la resistenza fisica del mantello (che verrebbe notevolmente ridotta dalla presenza di piccole quantità di materia fusa fra i granuli minerali o di bollicine interstiziali di gas).

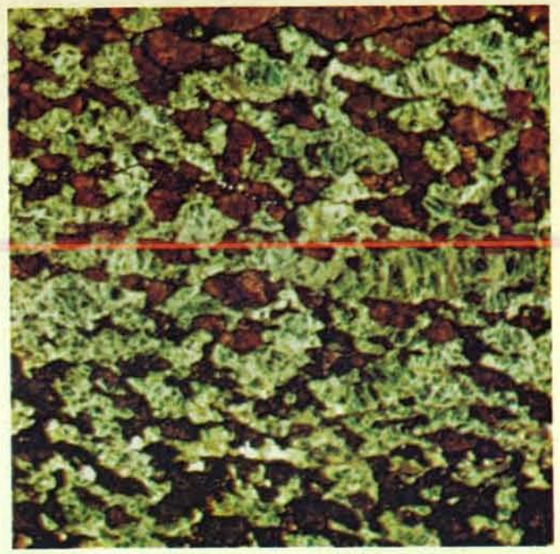
Acqua e anidride carbonica

In certi noduli peridotitici si trova della flogopite e dell'anfibolo, che sono minerali idrati, che, cioè, incorporano dell'acqua. Questa viene accettata come una prova dell'esistenza di acqua per lo meno in certe parti del mantello superiore. È dubbio se la concentrazione dell'acqua possa superare lo 0,1 per cento in peso, ed è probabile che la sua distribuzione non sia uniforme.

L'esame microscopico ha mostrato che i cristalli di olivina e pirosseno in alcuni noduli peridotitici di kimberliti e lave contengono numerose cavità dal diametro fino a 5 micrometri. Molte di esse sono piene di densa anidride carbonica liquida, intrappolata ad alta pressione. Ciò indica la presenza di anidride carbonica sotto qualche forma, nel mantello superiore.



Bollicine di anidride carbonica che compaiono in una micrografia elettronica di un campione di peridotite. L'ingrandimento è di 20 000 diametri. Le bollicine, che compaiono nelle discontinuità della struttura cristallina, dimostrano la presenza di anidride carbonica nel mantello. La micrografia è stata realizzata da Harry W. Green II.



La peridotite a mica e granato (*a sinistra*) e l'eclogite (*a destra*) sono rocce che provengono dal mantello. La peridotite consiste principalmente di olivina (*giallo verdastro*), ma contiene anche granato (*rosso*), pirosseno rombico e pirosseno monoclino (*verde brillante*). I campioni contenenti mica, un minerale che contiene acqua nella sua struttura, sono rari. Questo campione proviene dalla Tanzania. L'eclogite (questa proviene da una miniera in Sudafrica) è composta da granato (*rosso*) e da un pirosseno monoclino (*verde*). I minerali sono stratificati. Le rocce del mantello vengono portate alla superficie dai camini kimberlitici e da certe lave vulcaniche. Queste fotografie sono state fornite da J. B. Dawson, dell'Università di St. Andrews.

La microscopia elettronica ad alta tensione ha recentemente fornito dei quadri veramente dettagliati sui difetti reticolari, cioè sulle discontinuità nelle strutture cristalline che si trovano nei singoli granuli minerali. I difetti non possono essere visti al microscopio normale. Nelle micrografie elettroniche, lungo le discontinuità nell'olivina e nel pirosseno di certi noduli peridotitici, si vedono numerose bollicine di anidride carbonica. Ciò fa pensare che inizialmente l'anidride carbonica doveva essere disciolta nei minerali solidi, mentre poi per essoluzione è precipitata in goccioline di gas a causa delle deformazioni elastiche presso i difetti reticolari.

Le determinazioni indipendenti delle proprietà fisiche e chimiche del mantello devono essere coerenti fra di loro. Per controllare se sono coerenti bisogna conoscere le proprietà fisiche di un materiale di composizione uguale a quella stimata per il mantello nell'intervallo di pressioni e temperature che esistono in quest'ultimo. Se è possibile determinare in che modo varia la mineralogia in funzione della pressione e della temperatura si ha un modo per valutare la profondità da cui proviene un campione di roccia e la temperatura esistente a quella profondità quando la roccia vi si è formata o ha raggiunto l'equilibrio con l'ambiente circostante.

Pressione e temperatura

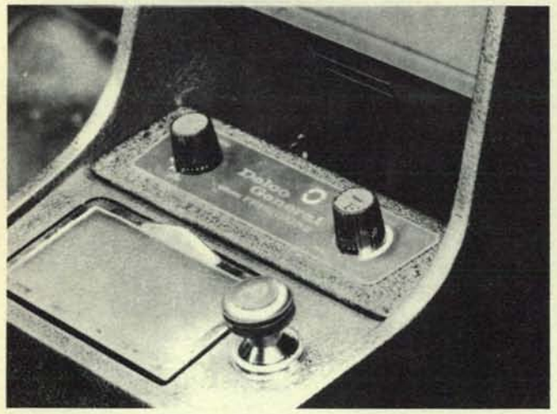
La struttura dei silicati è controllata dalle posizioni degli atomi di ossigeno; gli altri atomi, che sono molto più piccoli, occupano gli spazi fra di essi. A bassa pressione ogni atomo di silicio è circondato da quattro ossigeni, i centri dei quali formano un tetraedro; si dice che il silicio ha coordinazione quattro. A pressioni molto superiori gli atomi di ossigeno vengono schiacciati gli uni contro gli altri e si riorganizzano in una struttura più densa, con gli atomi di silicio a coordinazione sei. Questa riorganizzazione della struttura cristallina è una transizione di fase. Le variazioni brusche delle proprietà fisiche del mantello sono causate da transizioni di fase successive.

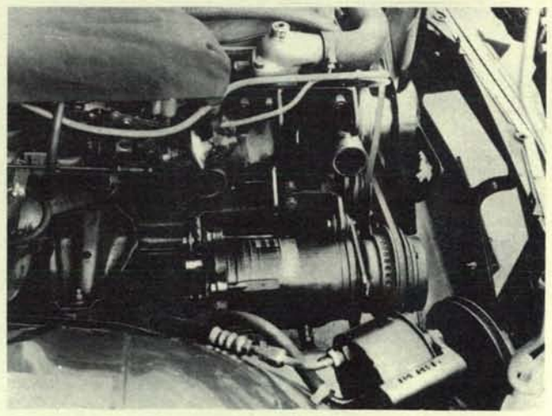
A partire da un materiale con una certa composizione, come la peridotite che si immagina che esista nel mantello, le transizioni di fase dipendono dalla pressione e dalla temperatura. Per via sperimentale è stato determinato un diagramma delle transizioni di fase della peridotite, con strumentazioni di laboratorio capaci di raggiungere i 200 kilobar (200 000 atmosfere, equivalenti alla pressione che esisiste 600 chilometri sotto la superficie terrestre). Questo diagramma, con metodi indiretti, è stato esteso anche a una larga fascia di pressioni superiori.

Dai noduli nelle kimberliti e nelle lave è noto che nel mantello le peridotiti possono cristallizzare almeno in tre associazioni mineralogiche: peridotiti a plagioclasio, peridotiti a spinello e peridotiti a granato. Gli esperimenti ad alta pressione dimostrano che queste associazioni sono legate da passaggi di fase. All'aumentare della pressione le peridotiti a plagioclasio si trasformano prima in peridotiti a spinello e quindi in peridotiti a granato (*si vedano le figure alle pagine 52 e 53*).

Alcuni studi sperimentali dimostrano che, a pressioni ancora superiori, la peridotite a granato subisce una transizione di fase che comporta un aumento di densità di almeno il 10 per cento; l'olivina, che predomina nel mantello superiore, si trasforma in un materiale simile allo spinello e il pirosseno alluminifero si trasforma in una struttura simile al granato, che si combina in soluzione solida col granato già presente. A pressioni prossime ai 200 kilobar i minerali vengono ancora compressi in strutture in cui tutti gli atomi di silicio hanno coordinazione sei, dando luogo a minerali sconosciuti alla superficie della Terra. Questa compressione risulta da un nuovo aumento di densità di circa il 10 per cento. La pressione a cui effettivamente avvengono i passaggi di fase aumenta all'aumentare della temperatura.

Ad ogni determinata profondità un





Nella foto a colori una vettura OPEL trasformata in una serra di freschezza da un condizionatore DELCO GENERAL.
Nella foto sopra, sono visibili le due manopole per regolare il flusso o la temperatura dell'aria.

L'impiego del compressore frigidaire, il nome più prestigioso nella produzione mondiale dei compressori, garantisce la qualità dell'impianto DELCO GENERAL.

aumento di temperatura può portare la roccia al punto dove inizia la fusione. Questa temperatura aumenta con la pressione, come è indicato dalla curva detta « solidus » nel diagramma delle fasi (*si veda la figura a pagina 52*). Una roccia composta da diversi minerali fonde progressivamente attraverso un intervallo di temperatura in cui cristalli solidi coesistono con del liquido. La fusione completa è contrassegnata dalla curva detta « liquidus ».

Gli effetti della temperatura si possono studiare per mezzo di una geoterma, cioè una curva che dà la temperatura ad ogni profondità all'interno della Terra. Se questa curva viene riportata sul diagramma di fase della peridotite ogni suo punto occupa una certa posizione sul diagramma delle fasi, e quindi definisce anche l'associazione mineralogica della peridotite a ogni profondità. Seguendo una geoterma sul diagramma delle fasi si può costruire una sezione trasversale attraverso un mantello ipotetico composto di peridotite. Ogni strato è costituito da una particolare associazione mineralogica.

I limiti fra gli strati del mantello si hanno presumibilmente alle profondità in cui la geoterma taglia i passaggi di fase. Risulta da tutto ciò che questi limiti corrispondono strettamente alle profondità alle quali cambia la velocità delle onde sismiche. Questo risultato viene considerato come una buona prova che la composizione del mantello è prossima a quella della ipotetica peridotite e che la struttura stratificata del mantello superiore è dovuta a transizioni di fase piuttosto che a cambiamenti di composizione.

La diminuzione nella velocità delle onde sismiche nella zona a bassa velocità si può spiegare con la presenza di acqua o anidride carbonica nel mantello superiore. Entrambe potrebbero provocare un inizio di fusione nella peridotite del mantello superiore. Il risultato è un cambiamento nelle proprietà fisiche della roccia. E se non fosse presente acqua, forse, un effetto simile potrebbe essere provocato da anidride carbonica intergranulare.

Nella prima metà di questo secolo la maggior parte dei geofisici era convinta che la convezione, cioè la risalita di materiale più caldo e la discesa di materiale più freddo, non poteva avvenire nel mantello, solido e rigido. Questa è una delle ragioni per cui per tanto tempo la teoria della deriva dei continenti non riuscì a trovare sostenitori. Recentemente, tuttavia, in un certo numero di modelli la convezione nel mantello è stata assunta come il meccanismo propulsore per la migrazione delle zolle di litosfera. I dettagli sul moto del mantello e sulla scala della convezione restano incerti, ma vi sono pochi dubbi sul fatto che le velocità sono estremamente basse – così basse che nella normale scala dei tempi umani il mantello è praticamente fermo.

Le zolle di litosfera, coi loro continenti sul dorso, si muovono alla velocità di qualche centimetro all'anno. Supponiamo che il materiale del mantello si muova alla velocità di cinque centimetri all'anno, il che equivale a circa 0,005 millimetri all'ora. La punta della lancetta di un normale orologio a pendolo da casa si muove a cinque centimetri all'ora, ed il suo moto non viene avvertito direttamente dall'occhio umano. Eppure questa velocità è 10 000 volte più alta di quella calcolata per il mantello. Ma anche così un moto di cinque centimetri all'anno è molto significativo nel tempo geologico; una particella di roccia potrebbe spostarsi dal fondo alla parte più alta del mantello in 58 milioni di anni, e cioè in una frazione molto piccola dei 4,6 miliardi che rappresentano l'età della Terra.

I moti del mantello

Come è possibile che le rocce scorrano, seppure così lentamente? Quando un fabbro prende una sbarra di acciaio freddo non è in grado di piegarla, ma se la riscalda fino a farla diventare rossa allora la può piegare facilmente, anche se la sbarra è ancora solida. Analogamente si possono deformare le rocce del mantello, che sono ad alta temperatura, anche se stanno ancora allo stato solido.

L'olivina, il pirosseno e le peridotiti sono stati sottoposti a deformazioni ad alta pressione e temperatura in una serie di esperimenti di laboratorio: si deformano davvero. Recentemente sono stati studiati al microscopio elettronico ad alta tensione i campioni deformati, per stabilire il meccanismo degli scorrimenti plastici. Una conclusione che si ricava da questo lavoro è che le deformazioni si verificano prima nei singoli cristalli, i quali poi ricristallizzano formando un mosaico di nuovi granuli. Il meccanismo dello scorrimento varia in funzione della temperatura e della pressione.

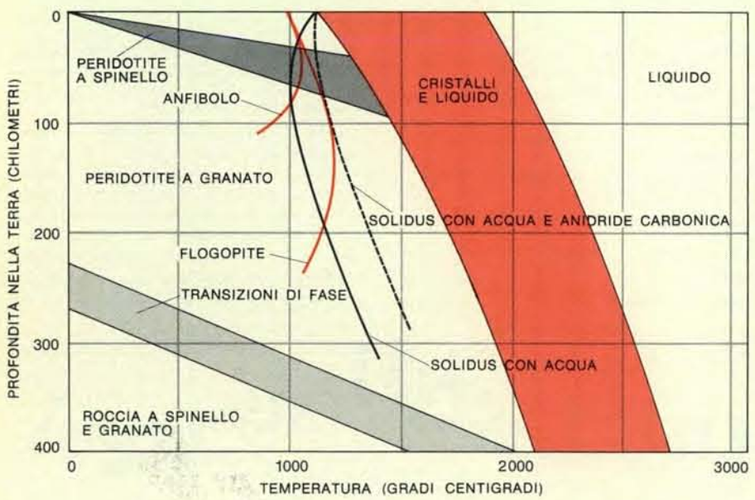
Sono stati proposti degli schemi di celle di convezione molto grandi, che si estendono per tutto il mantello (*si veda la figura nella pagina a fronte*). Un modello alternativo, che si basa sull'argomentazione che l'assorbimento di calore nelle transizioni di fase ridurrebbe la forza motrice, confina la convezione al mantello superiore, al di sopra della transizione olivina-spinello. Un terzo modello limita la convezione all'astenosfera, che è uno strato del mantello che si trova fra i 100 e i 300 chilometri di profondità. Tuttavia le zone di subduzione, dove il materiale della litosfera ritorna nel mantello e si collega quindi alle forze convettive, sembrano giungere fino a profondità di 700 chilometri. Un altro modello sul meccanismo che muove le zolle fa appello a « plumes » (ciuffi o pennacchi) termici nel mantello. Secondo questa teoria tutti i movimenti verso l'alto del materiale del mantello sono ristretti a circa 20 plume, ognuno dei quali ha un diametro di poche centinaia di chilometri e risale dal limite tra nucleo e mantello. Il movimento di ritorno viene effettuato tramite un lento moto verso il basso del resto del mantello.

Quando un plume raggiunge la litosfera il flusso diventa orizzontale, espandendosi radialmente in tutte le direzioni. Un plume crea un punto caldo, con attività vulcanica in superficie, e può causare una risalita della litosfera verso l'alto. In questo modo i plume potrebbero causare il movimento delle zolle di litosfera.

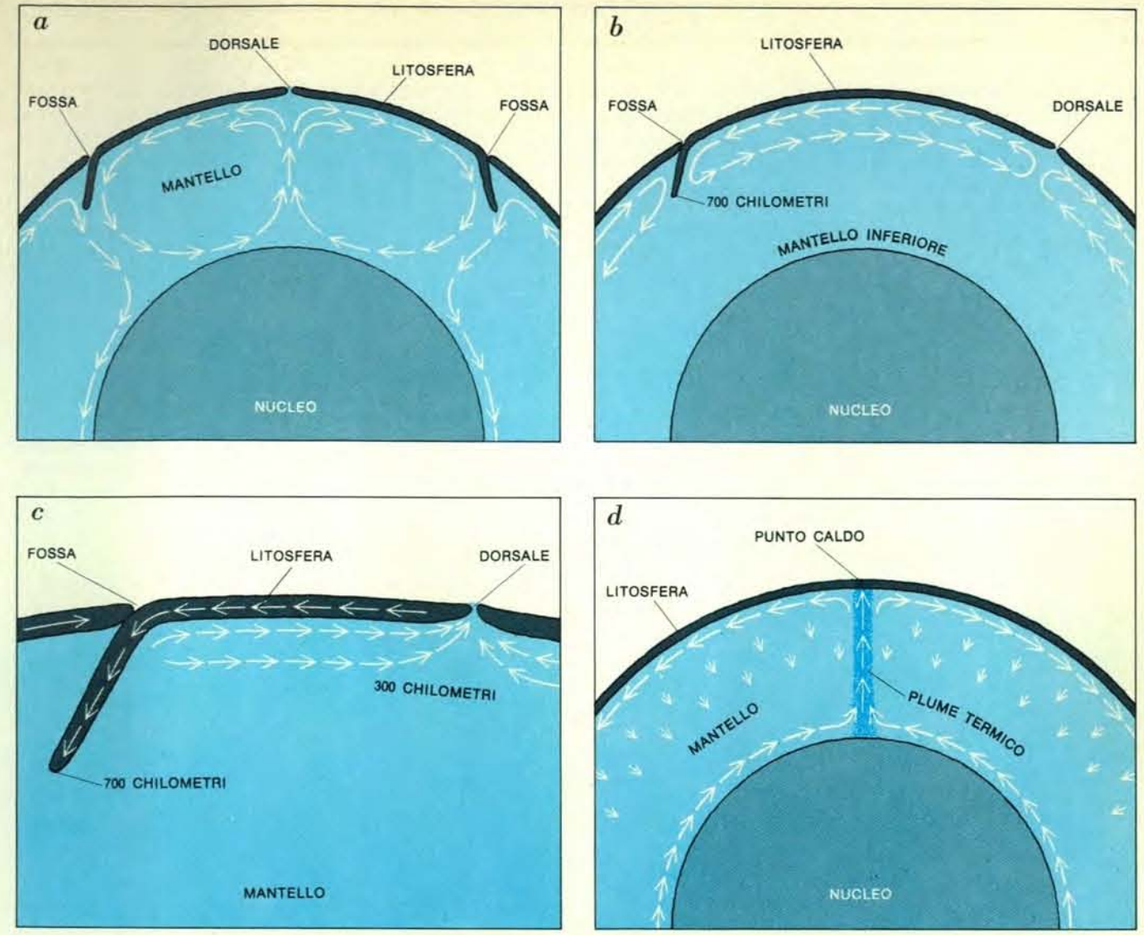
L'ipotesi del punto caldo attualmente è un'idea calda, e molti geologi ne stanno studiando le implicazioni per diversi fenomeni, fra cui le caratteristiche delle catene di isole vulcaniche come le Hawaii. Questa ipotesi, comunque, ora è stata messa sotto processo da parte degli studiosi della Terra che dubitano che un plume possa rimanere un tutto unico mentre risale per 2800 chilometri di mantello.

I moti verticali del materiale che costituisce il mantello, in qualsiasi tipo di sistema convettivo, portano dei cambiamenti nella distribuzione di temperatura; a una certa profondità la temperatura aumenta dove il materiale è in risalita e diminuisce dove sta affondando. Il moto cambia la forma della curva geotermica nel corso del tempo e da un posto all'altro a mano a mano che la convezione procede. La composizione dei minerali nella peridotite del mantello varia in funzione della pressione e della temperatura, e questo è stato stabilito dagli esperimenti di laboratorio sui diagrammi di fase delle peridotiti e dei minerali che la costituiscono.

Se una particella di peridotite del mantello, a causa dei moti convettivi in esso, viene sottoposta a variazioni di temperatura e pressione, la sua composizione mineralogica, per mezzo della ricristallizzazione, cerca di raggiungere un equilibrio con l'ambiente che sta cambiando. I movimenti sono così lenti che di solito l'equilibrio minera-



Lo strato a bassa velocità nel mantello superiore sembra che possa essere spiegato dalla presenza di acqua e anidride carbonica, che abbassano la curva del « solidus » e fanno fondere la roccia a temperatura e pressione più basse. L'anfibolo e la flogopite, minerali idrati, cioè contenenti acqua, sono stabili in un intervallo di temperatura limitato (*in colore*). Paragonando questo diagramma a quello per rocce del mantello asciutte si vede che la curva geoterma passa da una fase di peridotite solida a una di peridotite parzialmente fusa alla profondità di circa 100 chilometri, che corrisponde esattamente al tetto della zona a bassa velocità che si osserva nel mantello superiore.



Modelli di convezione proposti per spiegare come l'attività nel mantello muove le zolle di litosfera. Nella convezione il materiale più caldo si muove verso l'alto e il materiale più freddo si muove verso il basso. Secondo un modello (*a*) le celle convettive si estendono attraverso l'intero mantello. Nel secondo modello (*b*) esse sono confinate a profondità inferiori alla transizione spinello-olivina. Un terzo modello (*c*) limita i moti del mantello all'astenosfera. Nel modello del « plume » termico (*d*) tutti i movimenti verso l'alto sono limitati a pochi plume termici; quelli verso il basso sono estesi a tutto il mantello.

logico viene raggiunto. Se però la roccia viene improvvisamente trasportata alla superficie, come in un'eruzione kimberlitica, il campione non ha il tempo di riequilibrarsi e raggiungere la superficie col timbro mineralogico che corrisponde alla posizione e alla temperatura alle quali per l'ultima volta aveva raggiunto l'equilibrio nel mantello.

La composizione dei minerali coesistenti nelle peridotiti ad ogni pressione e temperatura è stata misurata direttamente con esperimenti di laboratorio. Usando questi dati come taratura oggi è possibile prendere dei campioni di peridotiti del mantello, come i noduli delle kimberliti, misurare la composizione dei minerali di ogni campione e valutare così la pressione (o la profondità) e la temperatura della zona del mantello da cui sono venuti.

Un'antica geoterma

Questi metodi, applicati ai noduli delle kimberliti, nell'ultimo anno o due hanno portato a risultati strani. Una serie di noduli, raccolti nello stesso camino kimberlitico, dà una serie di punti, uno per ogni nodulo, definiti dalla temperatura e dalla pressione di equilibrio stimate prima dell'eruzione. Il luogo di questi punti su un diagramma profondità (pressione) - temperatura corrisponde alla geoterma che esisteva al tempo dell'eruzione della kimberlite. In altre parole ogni nodulo della kimberlite porta immagazzinata nella sua composizione mineralogica la registrazione della sua temperatura e pressione d'equilibrio nel mantello prima di venire improvvisamente strappato via. I mineralogisti sperimentali impegnati in questo lavoro hanno trovato che i risultati dei camini kimberlitici del Sudafrica danno delle geoterme fossili a gradiente normale fino a circa 150 chilometri, ma con gradienti maggiori a profondità più alte; apparentemente a quelle profondità la temperatura era più alta del normale, per un certo intervallo di tempo prima che le kimberliti venissero eruttate, circa 100 milioni di anni fa.

Questi risultati sono estremamente interessanti per i geofisici che cercano di ricostruire la storia termica della Terra, la dinamica del mantello e le forze motrici della tettonica a zolle. Un'interpretazione di questi dati è che il punto in cui l'isoterma fossile si piega corrisponde al tetto dell'astenosfera a un'età di circa 120 milioni di anni fa, quando la zolla litosferica africana cominciò a muoversi rapidamente, mentre l'Atlantico si apriva. Secondo questa interpretazione il moto dell'astenosfera produceva energia termica per attrito e faceva spostare la geoterma verso temperature più alte e facendola piegare al limite litosfera-astenosfera, mentre la velocità di conduzione del calore attraverso le rocce è così bassa che la geoterma originale rimaneva praticamente invariata nella litosfera. I geofisici che concludono che l'attrito nell'astenosfera non è in grado di produrre effetti termici così vasti avanzano un'altra interpretazione. Essi argomentano che la parte più ripida della geoterma fossile, al di sotto del punto in cui si piega, potrebbe essere stata causata dalla convezione verso l'alto di un plume termico locale che avrebbe dato inizio all'eruzione della kimberlite.

Questo esempio illustra in che modo la tettonica a zolle, col suo centro nel mantello, ha riunito nelle stesse aule ai congressi scientifici ricercatori di campi che un tempo venivano considerati molto diversi gli uni dagli altri. Geologi di campagna, mineralogisti, geofisici, chimici e fisici sperimentali hanno insieme scoperto nei minerali dei noduli delle kimberliti una serie di informazioni che ora sono pane per i denti dei teorici che cercano di capire come il mantello si muove ora e come si è mosso nei 4,6 miliardi di anni di vita della Terra.

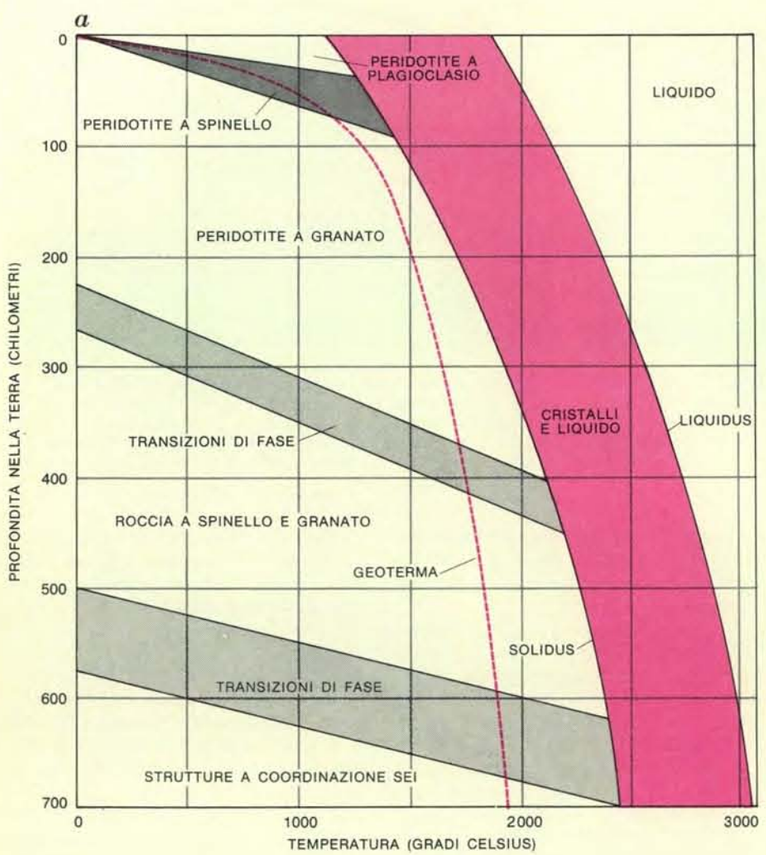
Vista da una nave spaziale, la Terra nel suo insieme si presenta come un grande globo rotante, liscio come una palla da biliardo. La più profonda perforazione eseguita alla sua superficie raggiunge una profondità di appena nove chilometri; non è che una puntura di spillo, che penetra per meno dello 0,15 per cento del raggio terrestre. Per questo è molto notevole che si sappia tanto sull'inaccessibile mantello.

Incertezza dei modelli

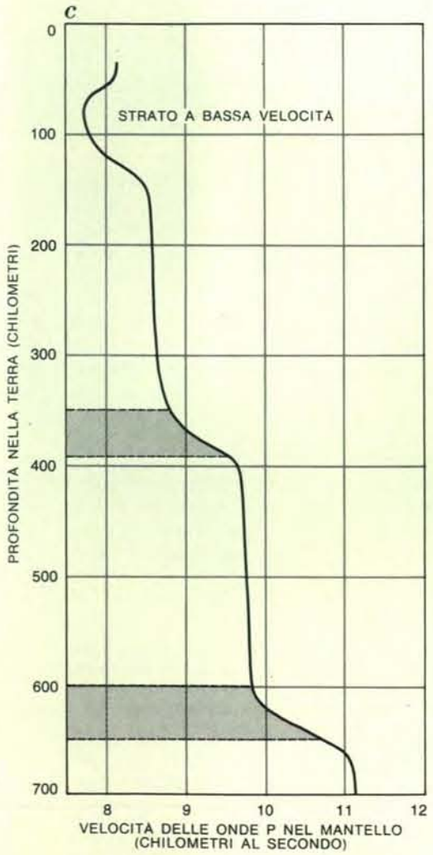
Ciononostante, la quantità delle informazioni nelle nostre mani è insufficiente per una piena comprensione del comportamento dinamico del mantello, ed è questa la chiave di molti fenomeni geofisici e geologici. La misura dell'incertezza su quello che realmente succede nel mantello si può illustrare paragonando due ipotesi sulla catena di isole vulcaniche delle Hawaii. Ogni isola si è formata in seguito a eruzioni al di sopra di una zona di fusione fissa nell'astenosfera. La zolla di litosfera, nel suo moto, ha poi allontanato l'isola, formando così, in un lungo intervallo di tempo, la catena di isole.

Secondo una interpretazione di questi eventi la fusione è localizzata in un punto caldo sopra un plume termico. Secondo un'altra interpretazione la fusione è localizzata per attrito in rocce che fluiscono dall'astenosfera in una colonna che si muove verso il basso attraverso il mantello. Si dice che questa colonna forma un'ancora gravitazionale, perché mantiene la zona di flusso verso il basso in una posizione più o meno direttamente soprastante.

Queste ipotesi diametralmente opposte vengono avanzate da studiosi di scienze della Terra assai rispettati. Le Hawaii vanno spiegate con un plume termico in risalita o con un'ancora gravitazionale che affonda? Io ho fiducia che presto l'accumularsi di nuovi dati e il raffinarsi delle teorie porranno dei limiti molto più precisi al quadro che si può fare della struttura e della dinamica del mantello.



Il diagramma delle fasi per la peridotite è stabilito in termini di pressione, o profondità, e temperatura. A una certa profondità un aumento di temperatura può portare la roccia al punto in cui inizia la fusione. L'inizio e la fine della fusione sono delimitati dalle curve del «solidus» e del «liquidus». Leggendo le due parti del diagramma dal basso verso l'alto, seguendo la geoterma (*a*), si vede che il mantello superiore ha una struttura stratificata a causa di successive transizioni di fase, dalla peridotite a spinello della parte più alta fino a una roccia composta di minerali con elementi a coordinazione sei, cioè con atomi di ossigeno molto ravvicinati che racchiudono atomi di magnesio, ferro e silicio, a una profondità di circa 600 chilometri. I numeri nella sezione trasversale (*b*) rappresentano le percentuali dei vari minerali che sono presenti nella roccia.



Il diagramma della velocità delle onde *P* mostra che le profondità effettive dei cambiamenti nel mantello, indicate dalle onde sismiche, coincidono strettamente con quelle delle transizioni di fase, ricavate sia dal diagramma delle fasi sia dalla geoterma riportati qui accanto.

INTRODUZIONE ALLA SESSUOLOGIA MEDICA
di **Giorgio Abraham** e **Willy Pasini.** Il primo vero manuale che offre al medico e allo psicologo una conoscenza approfondita ed esauriente del problema in tutti i suoi aspetti interdisciplinari. Un'opera unica per quanti sono interessati ai problemi pratici della sessuologia. Lire 8.000

IL COMPORTAMENTO SESSUALE UMANO
analizzato in sei culture rappresentative. A cura di **Donald S. Marshall** e **Robert C. Suggs.** Dalla ipocrisia repressiva all'estrema permissività. Un'analisi comparata delle molteplici e multiformi manifestazioni della sessualità promossa dall'Istituto Kinsey il più famoso centro di sessuologia del mondo. Lire 6.600

MEDICINA E POTERE
Collana diretta da G. A. Maccacaro

SCIENZA E POTERE
Un dibattito a più voci **(P. Bisogno, M. Cacciari, U. Cerroni, M. Cini, U. Curi, L. De Castris, G. A. Maccacaro, A. Massucco Costa, G. Napolitano, G. Quazza, G. Toraldo di Francia e altri)** per discutere e chiarire uno dei temi più attuali della società di oggi. Lire 2.000

ESSERE O MALESSERE
Le neurosi viscerali di **Massimo Gaglio.** Come la prevenzione e la cura del «malessere» trova nella scienza ufficiale che lo nega e nella società che tradizionalmente chiede la sua medicalizzazione, diffidenze e ostacoli. Un duro scontro tra la nuova medicina e il potere medico. Lire 5.400

2ª edizione
FARMACI SALUTE E PROFITTI IN ITALIA
di **Albano Del Favero** e **Giuseppe Loiacono.** Lire 3.200

Feltrinelli
in tutte le librerie

Passato e futuro dei centri storici

Il patrimonio culturale ed edilizio del nostro paese esige che stato, regioni e comuni intervengano con sollecitudine nel risanamento e nel restauro dei centri storici senza trascurare gli aspetti sociali ed economici del problema

di Luciano Pontuale

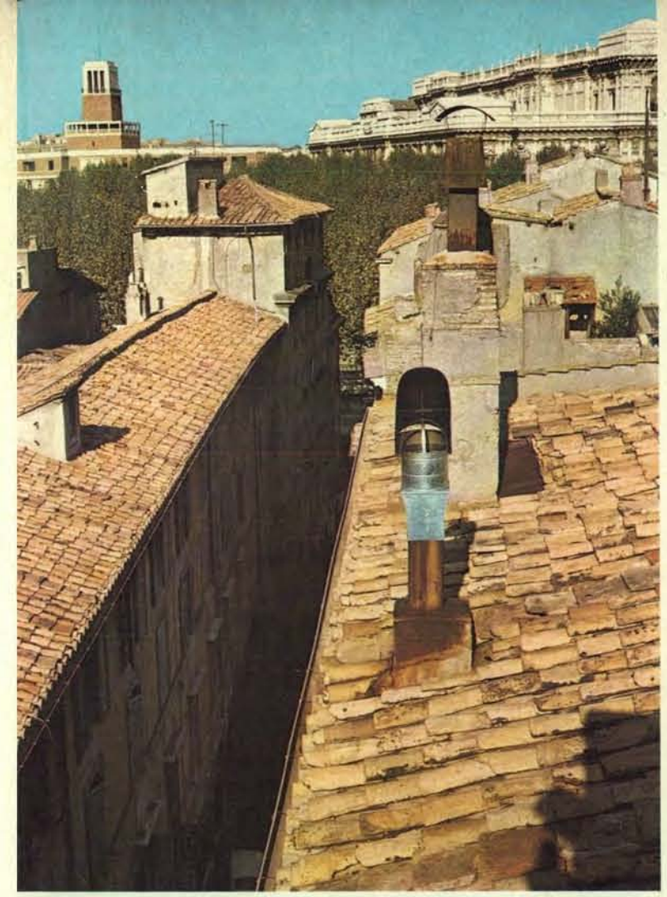
Tre anni or sono, la data del primo aprile 1972 ha costituito una svolta importante e, per certi aspetti, decisiva per una più diretta azione di conservazione e di recupero dei centri storici, a causa dell'avvenuto passaggio delle competenze in materia urbanistica alle regioni a statuto ordinario. Conseguentemente un diverso ruolo lo stato è stato chiamato a svolgere e nuove responsabilità sono state decentrate, con delega, agli enti locali.

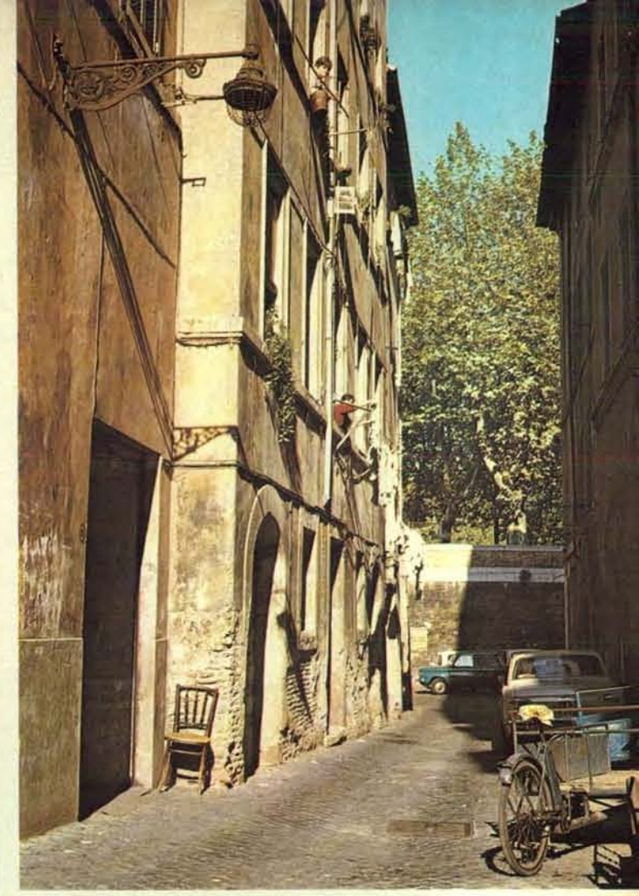
Tra stato e comuni un ruolo primario dovranno quindi svolgere le regioni. A queste ultime, in particolare, è stato assegnato, ormai da tre anni, stante la loro responsabilità nel settore urbanistico, il non facile compito di provvedere, attraverso una accorta e lungimirante azione politica di incentivi e di sostegno, alla predisposizione di programmi per la conservazione del patrimonio edilizio storico di oltre 8000 comuni italiani, guidando e sollecitando gli sforzi degli enti locali nella direzione ritenuta più idonea, sotto il profilo edilizio e urbanistico, onde soddisfare compiutamente le molteplici e complesse esigenze del risanamento edilizio, così come è richiesto anche dall'attuale cultura storica, economica e sociologica. Tutto ciò sulla base degli indirizzi che lo stato dovrà formulare per il coordinamento generale delle iniziative. Mancano ancora infatti i necessari inquadramenti, nonché i criteri di base per l'approntamento di tutte quelle metodologie operative che possano assicurare, a livello regionale o locale, l'attuazione di una più adeguata e unitaria disciplina dell'intera materia.

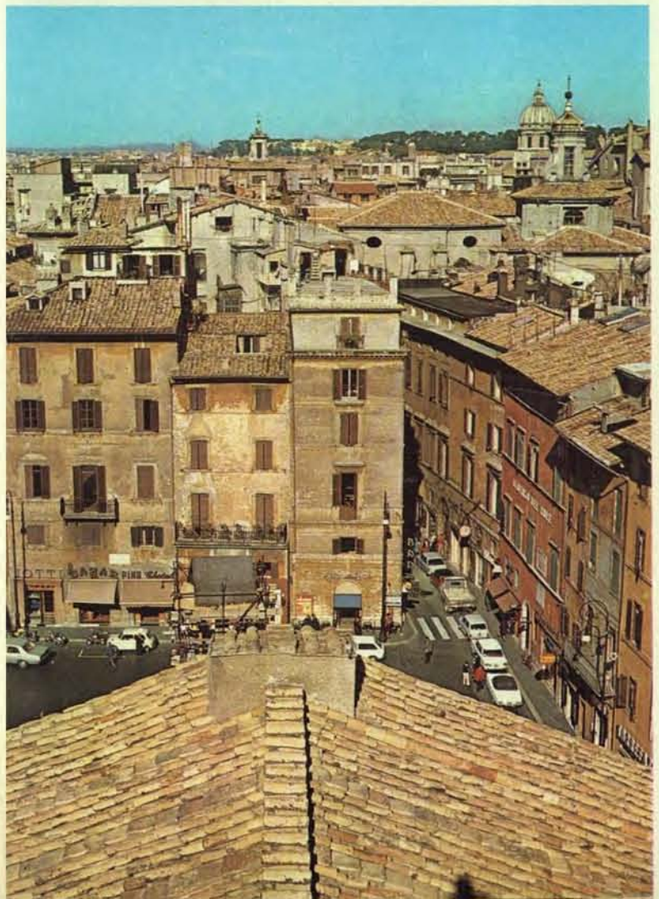
Alcune regioni, come per esempio la Toscana, l'Emilia Romagna e l'Umbria, tuttavia, anche in assenza di norme-quadro, si sono sollecitamente preoccupate della tutela del proprio patrimonio architettonico e hanno già predisposto numerosi seppur limitati provvedimenti legislativi o normativi. Questi, tuttavia, risentono ancora, come già detto, dell'assenza di un preciso inquadramento giuridico di livello generale che stabilisca, nel settore urbanistico, gli ambiti entro i quali l'intervento regionale può e deve articolarsi indipendentemente dagli aspetti ambientali, culturali o monumentali del problema. Occorre, in tale materia, non cadere nel facile equivoco, tuttora ricorrente, che per la tutela dei centri storici siano sufficienti solo quelle provvidenze legislative statali che solitamente vengono usate per i beni culturali di natura estetico-ambientale. I centri storici sono infatti strutture urbane che – indipendentemente dal costituire episodi storici che testimoniano valori di civiltà passate – contengono una pluralità di organismi di impianto complesso che possono essere recuperati solo privilegiando in essi gli aspetti più propriamente urbanistico-edilizi. Occorre, in sostanza, una precisa politica di intervento operativo che valuti le possibilità di riuso del patrimonio architettonico in termini strumentali, in ordine ai caratteri normativi, tipologici e, soprattutto, economici del problema.

È necessario quindi che le regioni, oggi, svolgano un ruolo autonomo, in qualità di « protagoniste », nel campo della promozione di concrete iniziative per l'attuazione sistematica di quei programmi di restauro e di risanamento conservativo che le più recenti risultanze propositive scaturite dal dibattito culturale dell'ultimo quindicennio hanno ormai chiaramente individuato nei loro termini sia individuali sia specifici e che lo stato si è anche di recente preoccupato di trasformare in concreti atti legislativi, anche se di carattere speciale, come nel caso di Venezia, Ancona, Tuscania, ecc.

Allo stato, alle regioni e agli enti locali sono, pertanto, oggi in molti a guardare con un certo interesse poiché spetta a loro attuare, nelle rispettive sedi, i diversi programmi di intervento. Lo stato di degradazione di circa 15 milioni di vani costituenti il patrimonio edilizio storico delle nostre città ha ormai raggiunto livelli tali da richiedere provvedimenti rapidi e risolutivi sia da parte dello stato – per quanto concerne la programmazione







La struttura rinascimentale di Roma, stretta nella penombra delle sue anguste vie (*in alto a sinistra*) si raccoglie ai lati dei lungotevere alberati al di là dei quali incombono gli edifici trionfalistici della città-capitale. In alcune zone in particolare, come nel quartiere Tor di Nona (*in alto a destra*), questa antica struttura ricca di edifici architettonici pregevoli ha subito in alcune sue parti profonde trasformazioni fisiche, sociali ed economiche a causa dell'isolamento a cui è stata costretta per la realizzazione dei lungotevere che ne hanno limitato la continuità. Come si vede nell'immagine, sullo sfondo, una antica via quattrocentesca è rimasta bloccata dal lungotevere sopraelevato che ha definitivamente eliminato il tradizionale accesso al fiume. A Roma, come in molte altre città italiane, che traggono la loro caratterizzazione dalle preesistenze di strutture di epoche diverse, non è infrequente la presenza di monumenti che risalgono a epoche molto antiche (per esempio, il tetto in primo piano nella fotografia in basso a sinistra raffigurante Piazza della Rotonda, è la copertura del pronao romano del Pantheon). Questo fenomeno, riscontrabile anche in Via della Rotonda (*in basso a destra*) ha caratterizzato ambienti inconfondibili nei quali convivono in simbiosi monumenti eccezionali ed edifici di scarso valore architettonico formando una unità ambientale che diviene essa stessa un unico monumento di eccezionale equilibrio umano. In ogni caso, i quartieri antichi rappresentano la testimonianza del passato civile della nostra società e, quindi, la loro conservazione si impone alla coscienza dell'uomo moderno in termini di sopravvivenza culturale e civile.

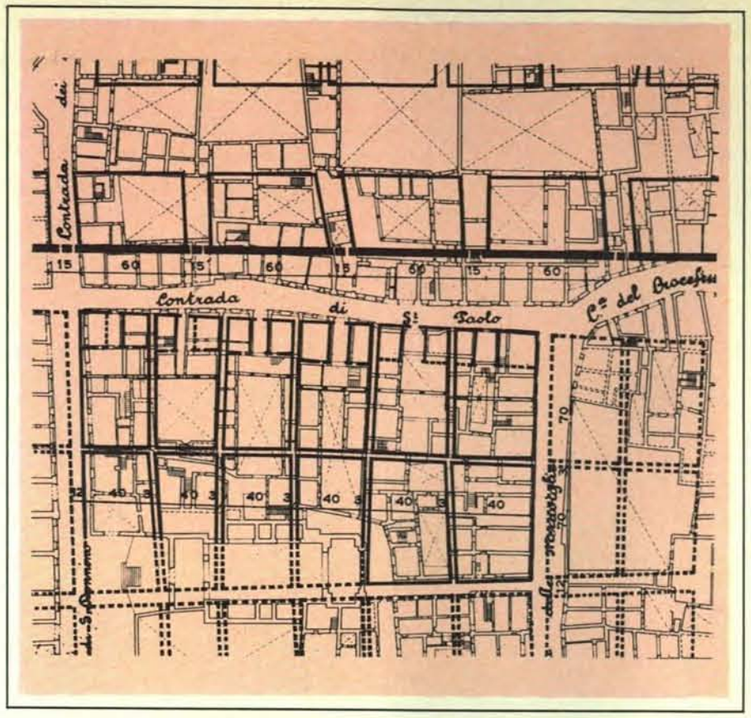
degli interventi – sia da parte delle regioni, per assicurare una attuazione unitaria dei programmi. Ai comuni spetta, per ultimo, il compito non facile ma determinante di provvedere alla gestione della intera operazione di risanamento con l'approntamento dei necessari strumenti attuativi.

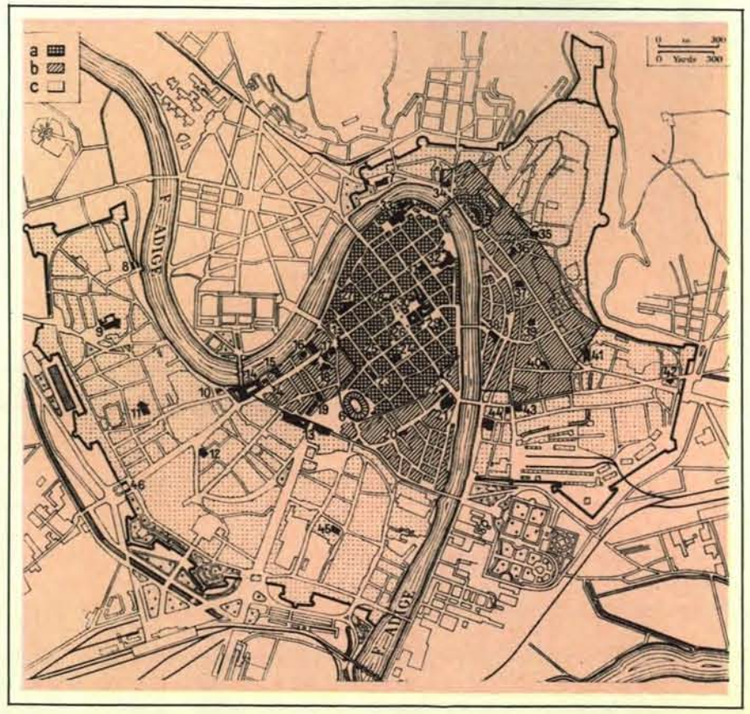
Tali provvedimenti dovranno mirare alla tutela dei caratteri più propriamente storici e tipologici degli organismi edilizi e determinare una loro più attenta riqualificazione sociale ed economica, così come è stato ormai sufficientemente acclarato da molteplici deliberazioni comunali.

Nessuno nutre più dubbi, ormai, che occorra evitare, sulla scorta della negativa esperienza dei decenni passati, il prevalere di tutti quegli interventi parziali di marca speculativa che tanti danni hanno arrecato al patrimonio edilizio e urbanistico dei centri storici, poiché in tal modo si verrebbe a incentivare ancor di più l'esodo di quelle fasce di popolazione tradizionalmente più deboli con conseguente modifica di quell'equilibrio sociale che è legato alla struttura economica degli abitanti.

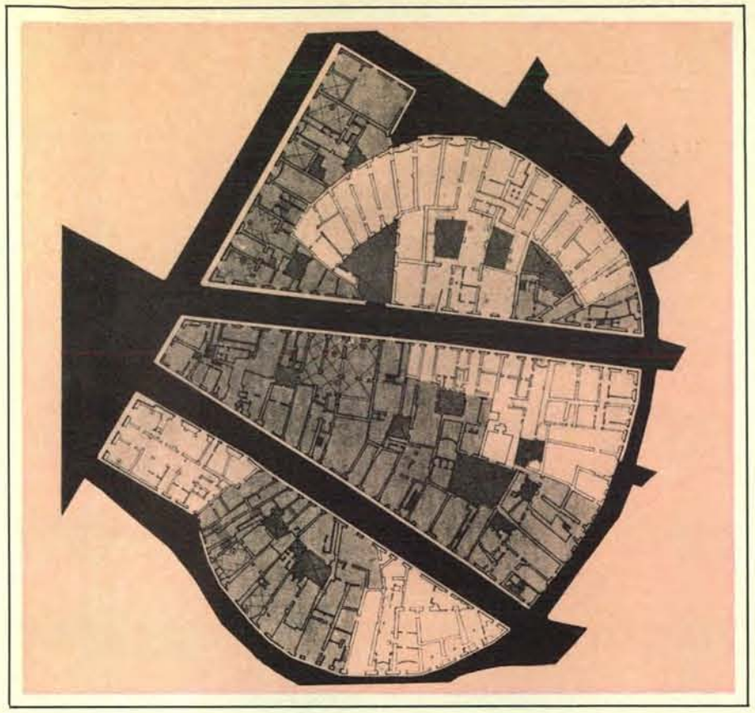
A differenza di quanto è avvenuto nel passato, occorre, pertanto, che il problema della riqualificazione dei centri storici venga definitivamente inquadrato in una prospettiva di sviluppo dell'intero territorio urbano in termini sia di razionalità degli insediamenti residenziali sia di pianificazione delle scelte territoriali, per una effettiva fattibilità delle proposte operative e per una loro maggiore concretezza a livello esecutivo. A tale proposito è bene ricordare che, se nell'immediato futuro sarà possibile giungere alla formulazione di più concrete iniziative – inquadrando cioè il problema del risanamento edilizio in un programma più generale e vasto di riequilibrio dell'intero territorio urbano – lo si dovrà soprattutto ai risultati metodologici e programmatici raggiunti nel corso degli anni sessanta. Tali anni, infatti, sono stati estremamente fecondi di idee e di contributi, il che ha fatto maturare con sufficiente rigoglio il complesso problema del recupero dei centri storici, anche se poi non v'è stato, in definitiva, l'emergere sia di un chiaro e costante indirizzo metodologico, sia di precise e unitarie proposte operative da contrapporre come valida alternativa all'intervento « spontaneo » che, tuttora, in gran parte prevale.

Il migliore contributo degli anni sessanta rimane, pertanto, quello circoscritto di una più generale e approfondita presa di coscienza del problema, da parte della classe dirigente, la quale

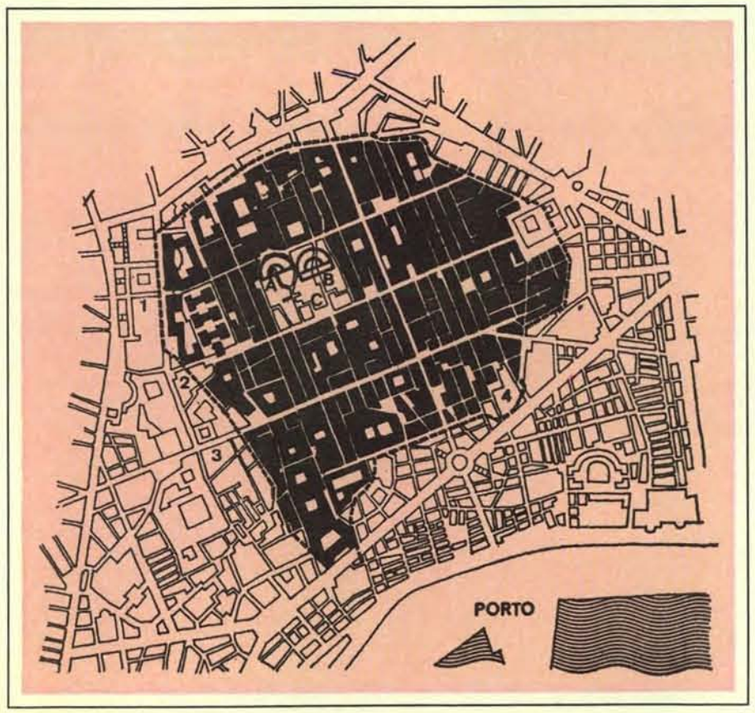




La presenza di strutture di epoca romana ha condizionato la forma architettonica e urbanistica di molte città italiane. La persistenza del *castrum* romano, per esempio, ha influenzato la configurazione di Como (*in alto a sinistra*), poiché costituisce tuttora la matrice degli attuali moduli spaziali entro i quali continuano a svilupparsi le abitazioni, mentre a Firenze (*in alto a destra*) i caratteri architettonici e urbanistici di molti quartieri, condizionati dalla presenza di strutture pubbliche romane, danno continuità, attraverso la permanenza degli elementi più antichi, alle espansioni urbane di epoca successiva. Tra le città romane, Verona, stretta nell'ansa dell'Adige e contor-



nata dalle espansioni medioevali (*in basso a sinistra*), racchiude entro le sue mura i valori tipici e inconfondibili della maggior parte dell'edilizia storica italiana, valori che, durante le manifestazioni indette in questo « Anno europeo del patrimonio architettonico », verranno portati all'attenzione di coloro che attraverso lo stato, le regioni e i comuni debbono assicurare la salvaguardia di ogni testimonianza delle civiltà passate. Per quanto riguarda Napoli (*in basso a destra*), inconfondibile appare ancora oggi la dimensione della città greco-romana, perfettamente conservata nelle strette strade costruite sul modello dell'isolato greco e sviluppatasi secondo gli schemi ippodonei.

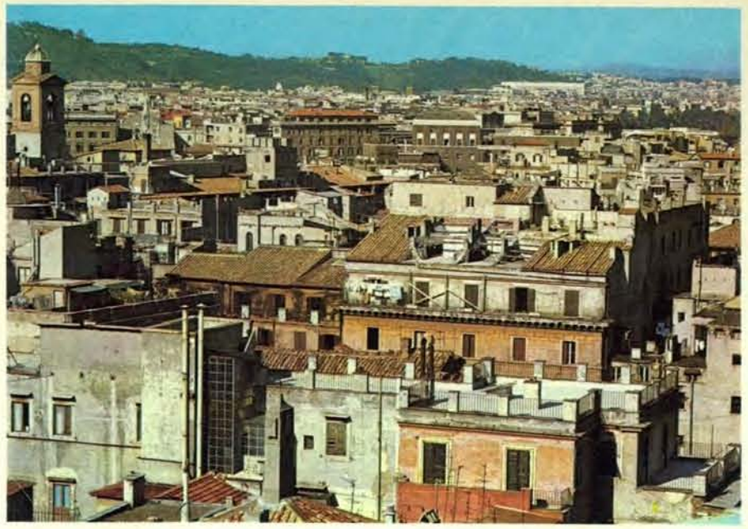


ha cercato spesso, in nome della cosiddetta « cultura ufficiale », di limitare attraverso i suoi organi tecnici le alterazioni che si andavano perpetrando ai danni del patrimonio storico. Ne è una riprova la maggiore attenzione che in quegli anni è stata riposta dagli organi statali di controllo in sede di valutazione delle scelte urbanistiche comunali, alle norme operative di Piano regolatore generale o di piano particolareggiato riguardanti i centri storici. (Il che ha comportato, di riflesso, anche una più accorta opera di pianificazione da parte degli stessi enti locali.) Nel fare un bilancio di tale periodo occorre sottolineare subito la particolare dimensione che è stata data al problema della tutela dei « beni culturali ». Gli anni sessanta si sono aperti con un importante convegno sul tema della « salvaguardia e risanamento dei centri storico-artistici » e si sono conclusi con la celebrazione dell'« Anno internazionale in difesa della natura ». Si è assistito, cioè, nel giro di dieci anni a una evoluzione qualitativa del concetto di « bene culturale », dai valori storici a quelli ecologici, e a una sua estensione quantitativa, dalla dimensione edilizia a quella territoriale.

Gli anni settanta hanno visto, invece, fin dall'inizio, lo spostamento dell'interesse degli studiosi più sugli aspetti di carattere economico caratterizzanti ogni bene culturale (specificatamente quelli edilizi) per una più adeguata opera di risanamento. Il che ha conseguentemente diminuito di molto l'importanza dei valori formali che erano alla base di tutte le discussioni del decennio precedente.

L'autunno del 1960 si chiuse con la dichiarazione finale del convegno dell'allora nascente « Associazione nazionale per i centri storico-artistici » la cui importanza fu tale da divenire, con il nome di « Carta di Gubbio », un documento fondamentale sulla salvaguardia più attiva dell'ambiente storico delle nostre città, i cui risultati avrebbero dovuto costituire il capitolo principale di quel tormentato « Codice dell'urbanistica » allora in corso di elaborazione.

I contenuti ideologici e programmatici che ispirarono l'azione rinnovatrice di quegli anni trovarono, tuttavia, la loro premessa e il loro principale sostegno proprio nel clima culturale, polemico e contradditorio del decennio precedente che vedeva l'Italia del « miracolo economico » sempre più tesa verso quel frenetico quanto caotico sviluppo edilizio che nel giro di pochi anni l'avrebbe portata alla crisi economica del 1963-1964. In quegli anni si



Dalla sommità della cupola del Pantheon si possono osservare le ampie trasformazioni cui sono assoggettati quasi tutti gli ultimi piani degli edifici della Roma storica che vengono spesso rinnovati per essere adattati come abitazioni di lusso. Ciò provoca, conseguentemente, profonde trasformazioni nella struttura sociale e abitativa dei quartieri antichi e tensioni economiche che facilitano l'espulsione dei ceti meno abbienti.

riteneva che fosse già una grande affermazione di principio il sostenere, per la tutela dei centri storici, la necessità della conservazione dei « volumi edilizi ». Era il periodo in cui i *mass media* facevano balenare modelli di vita basati sui consumi e le città si ampliavano sulla base di pochi, vecchi e inefficienti piani regolatori.

I piani di costruzione, quando c'erano, si limitavano a cristallizzare situazioni particolari senza tenere conto della necessità di assicurare una più efficiente strutturazione urbanistica delle città, mentre i nuovi piani regolatori, adottati o in corso di approvazione – che tuttavia avrebbero visto la luce solo nella seconda metà del decennio – erano strutturati sulla base della vecchia legge urbanistica del 1942, la quale era ben lontana da proporre qualsiasi pur minima azione di salvaguardia in difesa dei centri storici.

In presenza di tali e tante carenze nel settore dell'urbanistica italiana, in assenza di qualsiasi pur minimo ma razionale processo di pianificazione del territorio (sia a livello comunale sia intercomunale), le città erano quasi sempre inadeguate a sopportare lo sforzo edilizio del tempo.

Gli interessi immobiliari continuavano, pertanto, a concentrarsi nelle aree più centrali della città – cioè nelle zone di più sicuro profitto per la loro potenziale redditività anche se di più delicata struttura dal punto di vista storico-ambientale e monumentale – a causa della indubbia convenienza economica che aveva il proprietario per il maggior provento che gli sarebbe derivato attraverso il processo di trasformazione del proprio immobile mediante il meccanismo della demolizione e ricostruzione. Era la logica spietata di un certo sistema economico, frutto di un particolare calcolo di convenienza, che sotto l'alibi di un miglior « decoro » degli ambienti più « malsani » delle città tendeva invece a far aumentare il saggio di accumulazione nel settore edilizio anche attraverso l'aumento dei prezzi delle aree più centrali. In tali zone, molte abitazioni di elevato valore storico e ambientale, se non addirittura monumentale, per gli oneri derivanti dalle proprietà obsolete e insufficientemente ricompensati dagli affitti bloccati, subirono forzatamente questo progressivo processo di svalutazione e molti proprietari ridussero al minimo i costi di manutenzione; era nell'interesse esclusivo di un certo tipo di proprietà, infatti, in assenza di qualsiasi piano di risanamento e tenendo fissi i costi di esercizio, cercare di fare aumentare il più possibile gli introiti accrescendo l'indice di affollamento fino al giorno in cui si sarebbero presentate l'opportunità e la convenienza della sostituzione con nuovi edifici. Ciò al fine di ricavarne il massimo beneficio attraverso una operazione immobiliare che puntava il tutto per tutto sul sicuro richiamo che le nuove destinazioni d'uso (rappresentative, direzionali, commerciali, eccetera) nelle aree più centrali avrebbero certamente comportato a tutto beneficio dell'operazione stessa, secondo la logica speculativa tipica del settore. Nelle aree più interne, invece, ove ciò non poteva accadere per particolari vincoli dettati dalla ubicazione o dall'ambiente, si assisteva ugualmente al degrado delle abitazioni per un accentuarsi dei fenomeni di ghettificazione sociale delle aree stesse non sottoposte ad alcun processo di riqualificazione urbana ma lasciate alla progressiva incuria del tempo.

È il momento delle ultime episodiche trasformazioni edilizie che aggrediscono quasi tutti i principali centri storici, il cui peso diminuirà soltanto verso la fine degli anni cinquanta quando si renderanno ormai disponibili in numero sufficiente i primi piani regolatori del dopoguerra, piani che contenevano quanto bastava per assicurare il massimo tornaconto ai processi di sviluppo edilizio basati sulla preventiva acquisizione delle nuove aree fabbricabili.

Si venivano così, sostanzialmente, favorendo più o meno direttamente, con la connivenza di molte amministrazioni locali e a seconda dei calcoli di giochi politici legati al finanziamento delle reciproche « correnti » di partito, gli interessi edilizi di sempre più vasti gruppi monopolistici privati che cominciarono a valutare positivamente la convenienza al « decentramento » su quelle aree marginali alle città che loro stessi, sotto forma di società immobiliari, si erano nel frattempo affrettati ad acquistare a basso prezzo agricolo ma che poi i piani regolatori avrebbero « provvidenzialmente » trasformato in sicure aree fabbricabili: prendeva così slancio in tutta Italia quella vasta operazione edilizia che, nel giro di pochi anni, avrebbe distrutto il « giardino d'Europa ». In tale diverso clima edilizio, teso a esaltare la convenienza della « casa in proprietà » nei quartieri periferici, le aree più centrali cominciarono progressivamente a perdere di importanza; conseguentemente la pressione edilizia si allontanò sempre più sensibilmente da esse mentre l'accresciuta domanda di nuove case in proprietà (con tutti i benefici fiscali e le agevolazioni creditizie offerte dal governo) stimolava e favoriva l'esodo dei ceti borghesi e di quelli meno abbienti dai quartieri più antichi a tutto vantaggio di un nuovo tipo di operazione speculativa altrettanto redditizia quanto quella precedente, legata alla operazione di sostituzione edilizia. La possibilità di disporre di appartamenti vuoti da « restaurare », di preferenza attici, permetteva, sempre in assenza di precise scelte urbanistiche nel settore, le più grandi avventure immobiliari nel centro storico, che veniva progressivamente sottoposto, dietro la semplice conservazione delle sole facciate, a vaste operazioni di rinnovamento edilizio per quei ceti sociali di classe sempre più elevata che erano gli unici che potevano sopportare l'onere dell'impresa. Ancora una volta il problema si andava « spontaneamente » risolvendo in termini strettamente settoriali a solo vantaggio di poche classi privilegiate o di funzioni non residenziali contrastanti con il carattere dell'edilizia storica.

Gli ambienti culturali, tuttavia, reagirono vivacemente a tali manomissioni episodicamente impostate senza la garanzia di un qualsiasi pur minimo piano unitario e denunciarono all'opinione pubblica sempre più frequentemente i veri meccanismi speculativi che erano alla base delle principali alterazioni formali delle città.

Davanti a tale vasto movimento culturale che, trapassando il limitato settore dei semplici « addetti ai lavori », interessava sempre più ampi strati dell'ambiente artistico e letterario italiano con frequenti ripercussioni anche all'estero, gli organi pubblici, accusati pubblicamente di tanto scempio, avvertirono la necessità di una maggiore cautela nell'approvazione sia dei piani regolatori e dei programmi di fabbricazione sia dei progetti interessanti zone soggette a tutela paesaggistica o monumentale.

La più importante inversione di tendenza a livello ufficiale, al riguardo, venne proprio dal Consiglio superiore dei lavori pubblici nel 1952 quando – dietro la pressione della stampa e del mondo culturale più avanzato – tale consesso diede parere negativo per la prima volta a un tipo di operazione urbanistica che fino allora era stata invece sempre supinamente accettata ma che avrebbe comportato, se attuata, la devastazione di uno degli ambienti urbani più interessanti della Roma barocca. In quell'occasione venne infatti rifiutato il piano particolareggiato della via Vittoria che prevedeva l'ampliamento della strada in attuazione del Piano regolatore generale del 1931.

Da quel momento gli organi centrali di controllo in campo urbanistico cominciarono a percorrere una nuova più corretta strada a fianco di quelle stesse forze culturali che si erano fino ad allora battute, isolatamente, per la salvaguardia del patrimonio storico artistico.

Ma era necessario che passassero ancora 14 anni prima che il parlamento apportasse una sostanziale modifica alla vecchia legge urbanistica. La proposta Sullo di riforma urbanistica venne infatti clamorosamente bocciata nel 1963 ma la legge ponte del 6.8.1967, n. 765 – sulla scia dell'impressione provocata dall'inchiesta sullo scandalo edilizio di Agrigento in conseguenza della frana – si riuscì ad approvarla con le precisazioni dell'art. 17, quinto comma, che per la prima volta obbligava a conservazione e restauro tutti i centri storici la cui edilizia fosse anteriore al 1860, predisponendo, per la loro incolumità, rigidi sistemi di salvaguardia urbanistica in attesa dei piani particolareggiati divenuti da allora obbligatori.

È questo, tuttavia, il momento in cui la crisi dei centri storici farà sentire più che mai tutta la sua complessa dimensione. Allontanata quasi definitivamente la minaccia delle brutali sostituzioni edilizie, i nuclei antichi rimasero abbandonati a se stessi non essendo stata posta in cantiere nessuna iniziativa urbanistica per la loro effettiva riqualificazione come alternativa ai metodi allora correnti. Ben presto all'obsolescenza delle strutture si accompagnò una sempre più profonda decadenza sociale ed economica, il che aumentava le deficienze igieniche e rendeva precaria la permanenza in essi di sempre più larghi strati di popolazione.

I comuni, assillati da ben altri più pressanti e contingenti interessi, preferirono sistematicamente eludere il

---

Quest'anno si svolge in tutta l'Europa la celebrazione dell'« Anno europeo del patrimonio architettonico », proclamato tra tutti gli stati membri del Consiglio d'Europa, proprio per sottolineare la generale necessità di difendere ogni testimonianza urbanistica ed edilizia ancora presente nelle città storiche mediante provvedimenti atti ad assicurare la riqualificazione sociale ed economica nel costante rispetto dei valori culturali e ambientali tradizionali.

Anche in Italia è stato costituito un apposito Comitato nazionale posto sotto l'alto patronato del presidente della Repubblica, la cui presidenza è stata affidata al ministro dei lavori pubblici, Pietro Bucalossi, proprio per sottolineare l'importanza del problema edilizio e urbanistico che deve essere congiuntamente affrontato e risolto se si vuole far decollare una effettiva politica di recupero che ponga l'edilizia storica in condizioni tali da non subire, ulteriormente, i danni derivanti dal decadimento fisico, della degradazione socioeconomica o dalle trasformazioni connesse con una utilizzazione esclusivamente finalizzata al perseguimento di interessi speculativi.

In ordine a tali finalità, l'Italia ha ospitato a Bologna, nell'autunno scorso, il 2° Symposium del Consiglio d'Europa sul tema « La dimensione sociale della conservazione dei centri storici ».

Tale simposio ha sottolineato ancora una volta che occorre applicare in maniera rigorosa i principi del restauro urbano preservando gli aspetti sociali delle città e attuando ogni possibile intervento da parte dei pubblici poteri, sia per quanto concerne le provvidenze economiche sia per quanto riguarda l'informazione e la consultazione dei cittadini interessati al restauro dei propri alloggi.

In coincidenza con tale Anno europeo, nel 1975, il Comitato italiano ha promosso, attraverso il Ministero dei lavori pubblici di intesa con le regioni e i comuni interessati, una serie di studi e di ricerche sul tema dei centri storici avente come fine, tra l'altro, l'analisi e la comparazione dei seguenti problemi: aspetti della didattica universitaria; caratteri della normativa giuridica e legislativa sia statale sia regionale; contenuti metodologici espressi dagli strumenti urbanistici generali o attuativi; forme di intervento promosse dagli enti pubblici; ruolo dell'imprenditoria, della proprietà e dell'inquilinato; forme di partecipazione popolare connesse con il decentramento amministrativo e i problemi sociali; criteriologia degli interventi in atto nelle città di Ancona, Verona, Taranto e Bologna, prescelte come sedi di realizzazioni particolarmente significative in tema di risanamento conservativo.

Sono state, altresì, promosse varie iniziative per diffondere e divulgare i problemi dei centri storici nell'ambito delle scuole nonché per informare vaste categorie di utenti sull'importanza degli aspetti sociali, economici, politici, culturali, giuridici, tecnici attraverso sia programmi radiotelevisivi, sia concorsi giornalistici e fotografici, pubblicazioni e attività culturali, mostre, conferenze e dibattiti.

In autunno infine si svolgerà a Roma una « Conferenza nazionale dei centri storici » nel corso della quale saranno discussi i risultati degli studi e delle ricerche che verranno presentati dal Comitato italiano. Da tale conferenza dovrebbe scaturire un articolato normativo e un indirizzo propositivo da proporre all'attenzione delle autorità statali, regionali e comunali che debbono salvaguardare il patrimonio architettonico dei centri storici italiani.

problema con generici quanto comodi rinvii ai piani particolareggiati; si era nel momento in cui le amministrazioni comunali tentavano di impostare una nuova politica urbanistica per i territori esterni ai nuclei urbani sulla scorta delle legge 167 approvata nel 1962. Tale legge ebbe, peraltro, l'effetto di incentivare ancora una volta l'esodo della popolazione residente dalle aree più centrali, indirizzando larghi strati di abitanti verso i nuovi quartieri periferici con il miraggio, del tutto limitato, di migliori condizioni abitative.

Si crearono, così, i presupposti per una nuova più intensa domanda di abitazioni in zone ancora prive di valide infrastrutture urbane, mentre nessun incentivo veniva predisposto per un recupero qualificato di tutto quel vasto patrimonio edilizio esistente che era viceversa sufficientemente ricco di concrete attrezzature sociali, di effettive infrastrutture urbane e di qualificati servizi di quartiere, oltre che di una dimensione architettonica ed edilizia tutt'altro che squallida (se paragonata alla periferia) e di elevato valore umano e sociale.

Il 1967 – l'anno della legge ponte – trova ben pochi comuni in regola con gli strumenti urbanistici. Il loro numero non supera infatti il 10 per cento. A quel tempo pochissime erano le città dotate di un piano regolatore: Ferrara e Livorno si dotarono di un piano solo nel 1961, seguite l'anno successivo da Venezia e nel 1965 da Firenze e da Roma. Urbino, L'Aquila, Perugia e Fabriano ebbero approvati i loro piani solo nel 1967 ma, con l'entrata in vigore della legge ponte, quasi tutti gli strumenti urbanistici allora in vigore dovettero essere rivisti per adeguarli ai minimi stabiliti da nuovi standard urbanistici. In una realtà territoriale così difficile e complessa, il problema dei centri storici non poteva perciò minimamente essere affrontato e portato a soluzione dagli enti locali con la necessaria tempestività e comprensione. Ai comuni incombevano, infatti, ben pressanti problemi la cui contingenza, come già detto, scardinava qualsiasi programmata volontà di razionale indirizzo nell'azione pianificatrice. L'edilizia storica continuava, pertanto, a subire di riflesso tutti gli effetti negativi propri di tale situazione, anche se non è da sottovalutare la reale responsabilità che è propria di molte amministrazioni comunali nel campo dell'inefficienza urbanistica, da alcune volutamente accettata come presupposto per qualsiasi tipo di speculazione, o intrallazzo edilizio. Gli ambienti culturali, viceversa, erano quanto mai attenti all'evolversi della situazione e continuavano a dibattere l'argomento a qualsiasi livello: politico, tecnico, scientifico, sociale e di informazione.

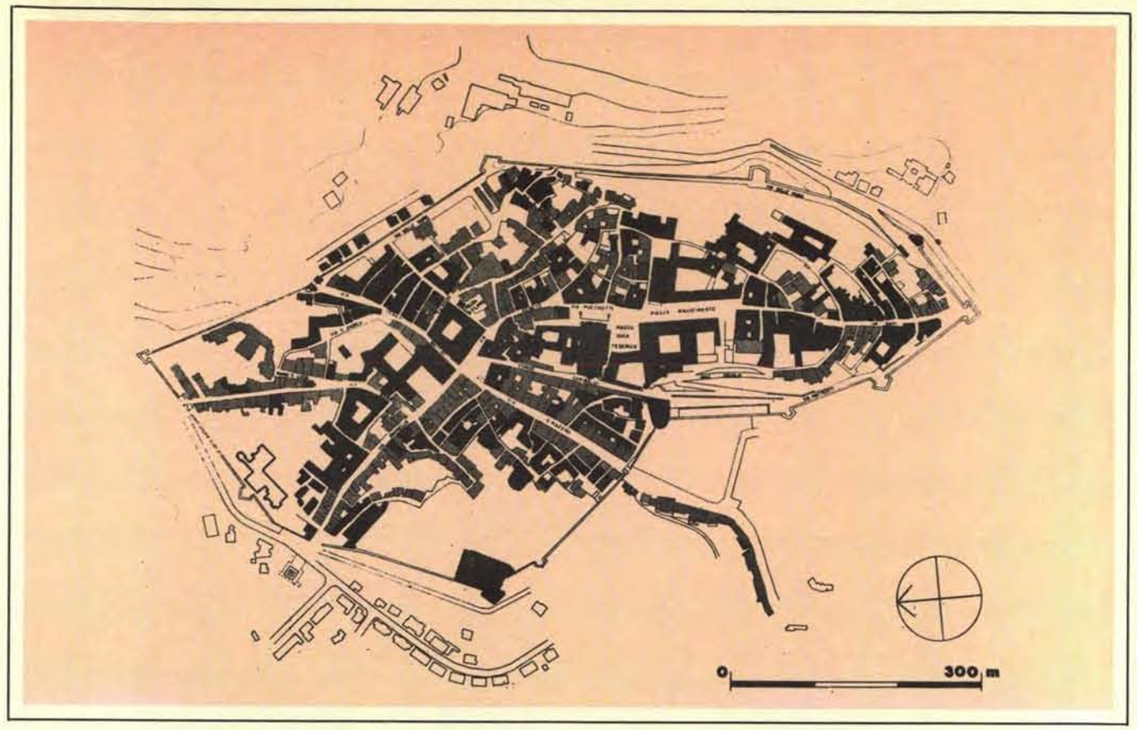
Il 1967 fu infatti anche il momento della mostra itinerante « Italia da salvare » inaugurata in quell'anno dal presidente della Repubblica a Milano, e promossa da Italia Nostra e dal TCI, la cui influenza sul pubblico e sulle autorità statali di controllo fu tale da esaltare una più attenta e consapevole presa di coscienza della necessità di tutela dei beni culturali. Gli anni sessanta videro, inoltre, attraverso il contributo dell'INARCH (Istituto nazionale di architettura) del 1963 (conferenza nazionale dell'edilizia) teso a sottolineare la necessità di una visione urbanistica e perciò più globale dell'intero problema e i contenuti della « Carta di Venezia » del 1964, promulgata dall'ICOMOS, il definitivo estendersi del concetto di « monumento » dalla singola opera architettonica all'intera città storica. Tale realtà culturale venne vieppiù sottolineata dagli incontri « antico-nuovo » di Venezia e Firenze (anni 1965-1966) che videro definitivamente dibattere, per quel decennio, il problema degli interventi architettonici nei contesti storici.

La Commissione Franceschini prima e la Commissione Papaldo dopo sottolinearono anch'esse (a livello parlamentare la prima e interministeriale la seconda) la necessità di pronte e urgenti riforme legislative per adeguare le vecchie strutture dello stato alle nuove necessità che i tempi ormai mutati richiedevano per una più adeguata tutela dei centri storici e dei beni culturali in generale. In conseguenza di ciò la classe politica, le amministrazioni pubbliche e gli enti locali, messi sotto accusa da una opinione pubblica sempre più larga e matura anche se ancora limitata nella cerchia ristretta ed elitaria di particolari classi sociali, sentivano il bisogno – rivelatosi poi del tutto illusorio – di fornire assicurazioni concrete circa la salvaguardia del patrimonio storico e culturale. Il comune di Urbino, affidando nel 1959 a Giancarlo De Carlo il Piano regolatore generale (PRG), fu il primo a rompere il cerchio di indifferenza nei confronti della tutela urbanistica dei beni culturali e sentì la necessità di promuovere una nuova politica di piano più rispettosa dei valori della città antica. De Carlo approfondì con particolare cura il tema giungendo a configurare una normativa di intervento estremamente dettagliata la quale, tuttavia, non sortì alcuno degli effetti sperati dall'urbanistica proprio perché le proposte, per quanto apprezzabili, non erano tali da calarsi compiutamente nella diversa realtà economica e sociale del comune.

L'insuccesso di Urbino, peraltro avvertito dallo stesso De Carlo fin dall'inizio, non costituì un fatto isolato; in tali anni verranno, infatti, scritte le pagine più sterili ed equivoche della nostra recente storia urbanistica.

Da una parte c'erano gli urbanisti i quali, basando dogmaticamente la loro operatività sopra una cultura accademica e professionale arretrata anche se mediata e perfezionata dalle ultime risultanze dei convegni, ritenevano che fosse sufficiente una buona metodologia operativa e una attenta normativa tecnica per risolvere compiutamente il problema dei centri storici; mentre dall'altra parte c'erano le amministrazioni comunali, che preferivano invece sistematicamente eludere l'argomento, certe che bastasse commissionare ai « tecnici » studi più accurati e approfonditi per accantonare temporaneamente il problema, per garantire la cultura e soddisfare l'opinione pubblica. Infatti con l'assicurazione del rinvio di qualsiasi intervento ai fantomatici « piani particolareggiati » che ben difficilmente sarebbero venuti alla luce, i comuni si creavano subito un prestigioso alibi culturale a sostegno delle loro « buone intenzioni » e della loro « inequivocabile volontà di portare a compimento con fermezza e chiarezza l'annoso problema del risanamento conservativo dei centri antichi... ».

In realtà era troppo evidente il profondo divario esistente tra le proposte operative degli urbanisti – che pur erano metodologicamente interessanti – e la realtà urbana, incapace di recepire una programmazione così complessa e delicata, in carenza di una precisa volontà politica per la tutela di tali beni e in mancanza di una effettiva presa di coscienza e di partecipazione al problema di strati sempre più vasti di popolazione. Le scelte culturali venivano così a scollarsi sempre di più dalla complessa realtà sociale ed economica dei centri storici e nessun contributo risultava pertanto fattibile sul piano pratico. Sono questi, infatti, gli anni fecondi di idee ma sterili di realizzazioni in cui Astengo progetta il risanamento del quartiere di S. Martino a Gubbio, il risanamento di Capodimonte e del Guasco ad Ancona (insieme a Coppa, Salmoni, e altri), il restauro di Assisi, di Bergamo e di Saluzzo; Coppa affronta i problemi dei centri storici di Perugia, Città di Castello, Orvieto e Vicenza; Piccinato quelli del ghetto di Padova e di Siena; Michelucci quello di Firenze; Pane quelli del centro antico di Napoli

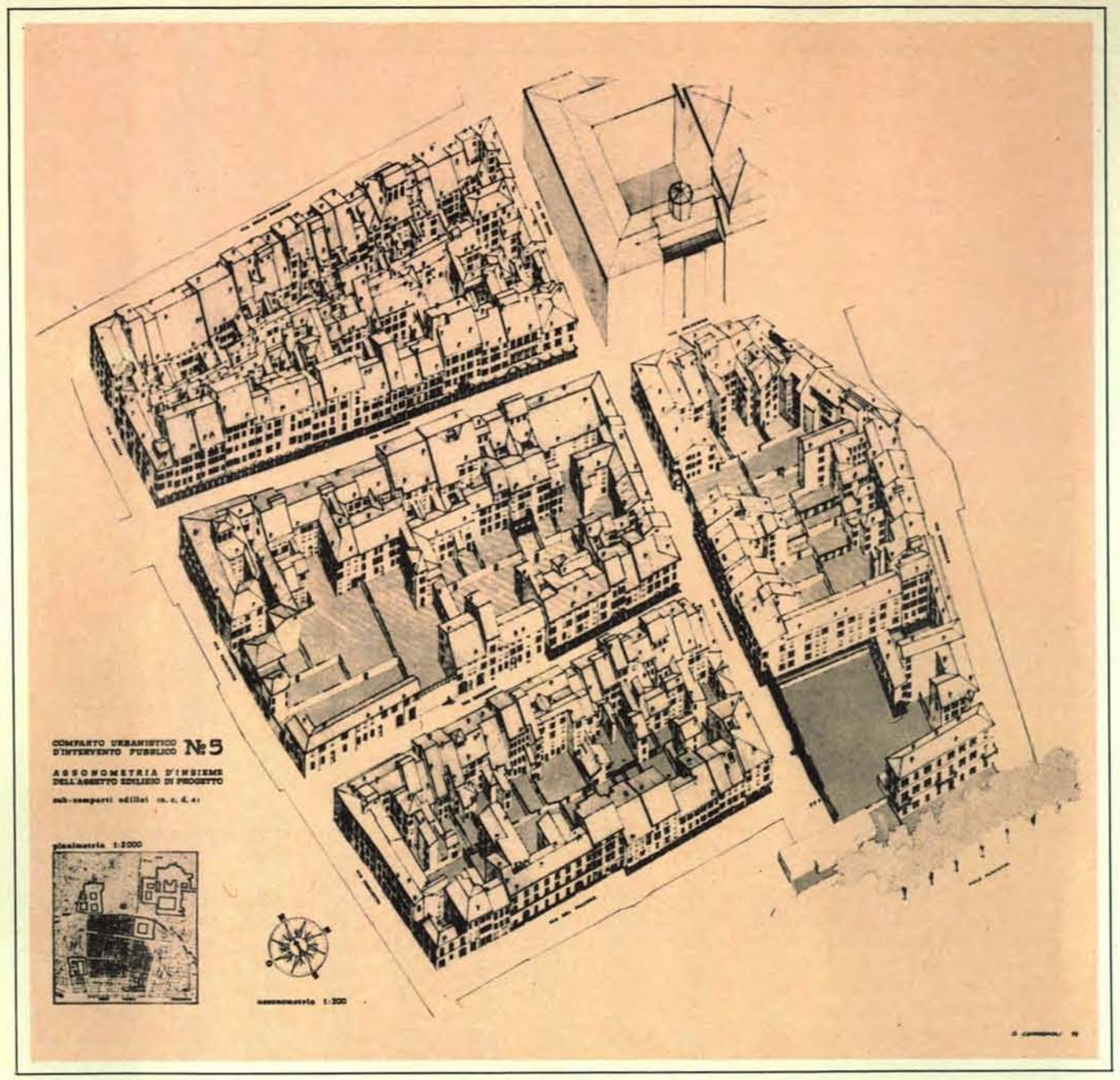


Il piano di risanamento del centro storico di Urbino, che appare in questa immagine, ha costituito una delle tappe fondamentali per promuovere una coscienza urbanistica volta al recupero della identità figurativa, formale e strutturale delle città antiche.

e di Molfetta; Quaroni quello di Bari vecchia; De Carlo, oltre a Urbino, quello di Rimini; Benevolo, insieme ad altri, quello di Ascoli Piceno e Ferrara; Ridolfi quello di Terni; Lugli quello di Tivoli e Forlì; Gazzola quello di Verona, mentre città come Bologna, Como, Reggio Emilia, Genova, Savona predispongono d'ufficio i rispettivi piani di risanamento; lo stato interviene pubblicamente con apposite leggi speciali sia per la conservazione di Bari vecchia, di Tuscania, del rione Terra di Pozzuoli, di Matera eccetera e sia per modificare più o meno profondamente le normative dei piani regolatori di città come Milano, Napoli, Roma, Venezia, Piacenza, Firenze, Savona, tanto per citarne alcune, al fine di assicurare una più rispettosa tutela dell'ambiente storico. Malgrado ciò il problema non è riuscito a trovare, nella realtà del nostro tempo, alcuna soluzione valida. I piani, le proposte, le prescrizioni e le leggi speciali, e non, sono rimasti e rimangono in gran parte sulla carta, inoperanti e inefficaci. Abbiamo a tutt'oggi una ricca messe di studi e di progetti ma ben pochi sono i risultati pratici che è stato possibile conseguire.

È evidente, pertanto, che fino a quando il problema della difesa dei beni culturali rimaneva una « obiettiva necessità » avvertita soltanto da quella piccola minoranza di persone che fino a ieri coincideva quasi solo con la cosiddetta « classe colta », era ben difficile arrivare a una soluzione efficace che portasse alla riqualificazione degli ambienti storico-urbani e che trovasse, a sostegno, il consenso delle popolazioni così com'era avvenuto, invece, per il problema dell'ecologia. Occorre, pertanto, oggi che nessuno pone più in dubbio la necessità della conservazione integrale delle città storiche in quanto tali, affrontare con diversa volontà politica (e qui sta il preciso nuovo ruolo delle regioni oltre che dello stato) la risoluzione dei problemi soprattutto d'ordine sociale ed economico che sono propri del settore, sia predisponendo la necessaria legge quadro urbanistica sia innescando un ampio processo di « partecipazione collettiva » alla gestione della città evitando, come invece avviene ancora in molte città, che l'interesse e le scelte della speculazione immobiliare continuino a modificare i contenuti residenziali tradizionali delle aree centrali, il che favorirebbe soltanto l'accentuarsi di quel vasto processo di sostituzione sociale che toglie spazio alla residenza, accelera definitivamente l'espulsione degli abitanti, modifica sostanzialmente il tessuto della città e altera i caratteri strutturali dell'edilizia storica.

Poiché esiste una serie di meccanismi economici relativi al tipo di sviluppo edilizio e urbanistico attuale che dice chiaramente no alla conservazione dei centri storici, sono in molti oggi coloro che ritengono che occorra aggredire proprio la politica della casa (le più recenti leggi sulla casa sono, ad esempio, meno discriminanti delle precedenti nei confronti delle abitazioni esistenti), facilitando il recupero dell'edilizia storica a vantaggio soprattutto di quelle classi sociali, medie e piccole, che sono state in questi ultimi anni forzatamente espulse dai loro luoghi tradizionali sia per l'alto costo delle operazioni di risanamento sia per il calcolo speculativo legato alle scelte della grande proprietà privata o per gli squilibri territoriali tuttora esistenti tra aree depresse e zone di accentuato sviluppo industriale. È paradossale che in Italia, paese ove esiste tuttora un notevole fabbisogno di abitazioni, si lascino decadere circa 15 milioni di vani antichi o se ne modifichino i contenuti tipologici per adibirli a destinazioni di uso diverso dalla residenza tradizionale. Occorre, in definitiva, che tali beni – che costituiscono



Gli interventi di restauro, riguardanti il centro storico di Bologna ed esemplarmente condotti attraverso l'analisi delle tipologie edilizie, hanno permesso soluzioni estremamente rigorose per ogni testimonianza storica del passato, pur tenendo nel massimo conto le esigenze dell'intervento pubblico e della tutela degli interessi economici e sociali delle popolazioni esistenti.

un insieme culturale ed economico considerevole – vengano conservati, indipendentemente dai loro valori formali (belli o brutti che siano), oltre che per i loro valori storici, anche per i contenuti sociali. Questi ultimi potrebbero essere convenientemente recuperati solo se si riuscisse a garantire una diversa impostazione economica del problema. Per far ciò occorrono indubbiamente provvidenze politiche e legislative che non siano di solo « vincolo » bensì tali da incentivare un risanamento effettivo delle abitazioni per un recupero di sempre più larghi strati sociali (e a vantaggio soprattutto della popolazione più giovane) che altrimenti tenderebbero a spostarsi verso abitazioni, forse apparentemente più confortevoli, ma certamente poste in zone sempre più periferiche e marginali.

Ciò potrà risultare fattibile solo se le regioni sapranno impostare per i rispettivi territori una loro politica di intervento che modifichi il modello di assetto territoriale in atto come alternativa all'attuale modello di concentrazione per un più equilibrato uso delle risorse e del suolo. È necessario cioè, prima di affrontare i problemi più specifici del « risanamento », cercare di risolvere l'attuale squilibrio territoriale tra Nord e Sud, tra zone di sviluppo e aree in abbandono. Ma per far questo occorre soprattutto una diversa politica di localizzazione industriale, una nuova riqualificante azione a sostegno dell'agricoltura e, a livello di crescita urbana, l'attuazione di tutta una serie di incentivi e di disincentivi per evitare lo squilibrio nelle previsioni urbanistiche di quasi tutti i nostri comuni dediti fino a ieri troppo semplicisticamente alla decuplicazione delle loro effettive necessità di fabbisogno abitativo.

Le provvidenze economiche previste nella legge speciale per Venezia dovrebbero, in questo quadro, tendere a sollecitare e a facilitare il rientro nella città insulare di gran parte della popolazione che forzatamente è stata costretta ad allontanarsi da essa in questi ultimi anni.

Ed è proprio per ottenere ciò che la legge prevede la realizzazione di un « piano comprensoriale », come strumento volto alla riorganizzazione del territorio circostante la laguna, se si vuole una effettiva salvaguardia e rivitalizzazione di Venezia. Solo con un contributo della spesa pubblica sarà tuttavia possibile raggiungere risultati tali da ridurre al minimo i processi degenerativi, frutto dell'attuale speculazione edilizia e fondiaria.

Analogo è il caso del centro storico di Napoli vincolato per il suo elevato valore storico dal Consiglio superiore dei lavori pubblici a restauro e risanamento conservativo, con una rigorosa normativa urbanistica, che tuttavia non potrà attuarsi senza una precisa partecipazione dello stato a causa dell'elevato costo delle operazioni di restauro e del carattere stesso dell'opera di risanamento, la quale non può essere lasciata alla sola iniziativa privata.

Occorre ridurre al minimo gli incentivi allo sfruttamento incontrollato della consistenza edilizia delle città storiche, come tuttora invece si lascia fare grazie anche all'attuale meccanismo di risanamento che, come nel caso di molte città, dovrebbe godere di differenziate condizioni di rimborso e non di semplici contributi e incentivi. Solo fermando l'esodo delle popolazioni, migliorandone le classi di età oggi poco elevate, eliminando i processi di ghettificazione, sarà possibile rovesciare il letale fenomeno di obsolescenza delle strutture urbane storiche. Questo è, indubbiamente, il problema più difficile che le regioni – a livello politico – dovranno affrontare per un immediato riequilibrio dei valori culturali su tutto il territorio, al fine di proporre soluzioni realisticamente valide e non ipotesi disancorate dalla realtà come di fatto finora purtroppo è accaduto.

Nel settore, invece, più specificatamente teorico-propositivo e metodologico le regioni possono disporre oggi di soluzioni operative quanto mai valide sotto il profilo dei possibili risultati conseguibili, grazie al passaggio avvenuto in questi ultimi anni da una concezione meramente figurativa e formale, cioè estetizzante, a una concezione più rispettosa dei valori storici di ciascun organismo edilizio visto nell'integrità della sua concezione tipologica, formale e strutturale.

Il Convegno di Bergamo del 1971 e gli studi dei piani di Bologna, Como e Rimini, oltre a quello di Vicenza fatto da Coppa, hanno per la prima volta permesso di dare una nuova impostazione al problema dei centri storici.

Per quanto riguarda gli interventi operativi le più recenti esperienze hanno, infatti, dimostrato che i caratteri tipologici (e quelli più strettamente a questi connessi come quelli strutturali e distributivi) costituiscono, nell'ambito del loro lento ma coordinato sviluppo temporale, in stretta simbiosi con gli stessi valori stilistici, gli elementi architettonici caratterizzanti la complessa unità e omogeneità di qualsiasi organismo edilizio storico. Poiché occorre tendere alla conservazione e alla riqualificazione dei caratteri costitutivi dell'intero organismo edilizio, così come gli eventi storici lo hanno configurato fino ai nostri giorni, è necessario che non vengano mai disgiunti gli aspetti formali da quelli tipologici se si vuole compiere effettiva operazione di restauro edilizio, nel recupero cioè di quei caratteri strutturali, distributivi e di destinazione di uso che, unitariamente considerati, costituiscono gli elementi essenziali per la effettiva conservazione della complessa realtà storica e sociale di ogni città.

Tale metodologia dovrà, inoltre, essere sempre verificata con l'utenza al fine di una maggiore, più precisa, puntuale e pertinente rispondenza delle necessità del risanamento ai bisogni concreti della popolazione.

La nuova « Carta del restauro » emessa di recente dal Consiglio superiore delle belle arti e quella in corso di formulazione da parte dell'ICOMOS per aggiornare la Carta di Venezia del 1964 costituiscono, in proposito, il più importante e significativo contributo « ufficiale » al problema. In tali documenti, infatti, i centri storici vengono considerati per la prima volta meritevoli di restauro e risanamento consecutivo non solo per i loro valori estetico-formali e per i loro caratteri storico-documentari, ma soprattutto per la presenza di quegli aspetti funzionali, distributivi, strutturali che sono tali da determinare i « caratteri tipologici di ogni singolo organismo edilizio » e che sono estremamente importanti da conservare per garantire una effettiva unità all'operazione del restauro urbano.

Tale atteggiamento « ufficiale », unitamente alla nuova, più dilatata dimensione scaturita dal mondo della culture, fa sì che oggi per risolvere compiutamente ogni aspetto del complesso problema non possano più venire ignorati anche quei meccanismi economici che sono stati finora i maggiori responsabili delle più pesanti trasformazioni dei centri storici.

Una efficace politica di riequilibrio di tali centri, pertanto, pur partendo dalle ultime rigorose metodologie operative valide a livello urbano ed edilizio, dalle nuove definizioni contenute nella « Carta del restauro », dalle precise scelte politiche ed economiche formulate dagli ultimi congressi in materia, non potrà mai essere formulata pienamente se non si prenderanno in esame, come già detto, le complesse implicazioni territoriali – e quindi urbanistiche – che hanno portato alla grave crisi di questo importante settore, il cui riequilibrio è fondamentale se si vuole riportare ordine nella città. Il recupero dei centri storici è, infatti, strettamente connesso a un effettivo riassetto del territorio proprio per un migliore utilizzo di ogni sua risorsa, dal momento che esiste una stretta dipendenza tra i meccanismi di sviluppo della città, le variazioni del mercato immobiliare e le scelte della proprietà sul capitale investito nel settore edilizio. È da questi problemi, infatti, che nascono poi conseguentemente le scelte per le operazioni di restauro e di riqualificazione sociale o non. È necessario giungere quindi al più presto a una legge quadro per l'urbanistica che contenga un apposito titolo sui centri storici, il quale tenga nel giusto conto il problema della conservazione dell'edilizia di valore storico, il tutto nell'ambito di una sostanziale revisione delle attuali politiche di sviluppo del territorio e delle città, alla luce delle ultime acquisizioni del mondo culturale, economico e sociale in materia di risanamento conservativo e della recente esperienza legata all'applicazione della legge speciale per Venezia.

Il 1975 è l'anno internazionale per la conservazione dei centri storici promosso dal Consiglio d'Europa similmente all'anno internazionale per la difesa della natura del 1970; saranno capaci lo stato, le regioni e i comuni, in occasione di questo appuntamento europeo, di mostrare tangibilmente che esiste una volontà effettiva per risolvere il problema, o ci si ridurrà ancora una volta a vaghe quanto generiche affermazioni di principio? Gli ultimi venti anni di storia europea non lascerebbero, in proposito, sperare nulla di buono per il futuro ma oggi, dal momento che in Italia è stata profondamente modificata la struttura dello stato con l'avvento delle regioni, è lecito, almeno, sperare in un ripensamento. Rimarrà anche questo solo una speranza?

Orologi biologici della zona di marea

Negli organismi sono assai frequenti orologi interni regolati sul ritmo del giorno solare; piante e animali che vivono lungo la costa possiedono però anche un orologio basato sul giorno lunare

di John D. Palmer

Nella spiaggia situata tra le linee di marea sulla costa settentrionale di Cape Cod vive un'alga bruno-dorata microscopica, la diatomea *Hantzschia virgata*. Il protoplasma di questa pianta unicellulare è racchiuso in una parete cellulare silicea, allungata e perforata in molti punti da pori e fessure. Attraverso alcuni pori terminali l'alga secerne una sostanza mucillagginosa che utilizza per compiere un lento movimento di propulsione attraverso il suo habitat sotterraneo. Durante il periodo di bassa marea diurna il minuscolo organismo mobile scivola tra gli interstizi dei granelli di sabbia fino alla superficie: qui rimane per tutto il tempo di bassa marea, esponendo il suo apparato fotosintetico alla luce solare. A metà estate le diatomee sono talmente numerose che, nonostante le dimensioni microscopiche, formano un evidente tappeto di color bruno-dorato che riveste la spiaggia. Esse, non appena l'alta marea le sommerge, si sprofondano nel sicuro rifugio della sabbia.

Per mettere in evidenza alcuni interessanti aspetti del meccanismo migratorio di questi microrganismi si trasporta la sabbia contenente le diatomee dalla costa settentrionale di Cape Cod al Laboratorio di Biologia marina di Woods Hole, sulla costa meridionale. I campioni vengono messi in un termostato, dove la temperatura è mantenuta costante e la luce rimane accesa in continuità: in questo nuovo ambiente, in cui non esistono i cambiamenti periodici giorno-notte e alta-bassa marea, le diatomee continuano le escursioni periodiche fino alla superficie della sabbia, in sincronia virtuale con le diatomee rimaste a 45 chilometri di distanza. I movimenti nel laboratorio sono talmente « puntuali », che quando progettiamo un'escursione a Cape Cod Bay, talvolta osserviamo le diatomee nel termostato anziché consultare le tavole di marea. Poiché il ritmo delle diatomee persiste in assenza delle periodicità ambientali che dovrebbero regolare questo comportamento, sembra che all'interno di queste piante vi sia un orologio biologico che regola le caratteristiche temporali della loro vita.

Questa affermazione non è uno dei tanti divertenti aneddoti che si raccontano sulla vita degli organismi: la maggioranza degli abitatori della zona di marea mostra infatti a precise scadenze ritmi periodici, caratterizzati dalla ripetizione di un fatto fisiologico o relativo al comportamento, come una febbrile attività sincronizzata con una particolare fase della marea. Poiché in ogni giorno lunare (lungo 24,8 ore, che rappresentano il periodo di tempo che intercorre tra due successivi levar della luna) si producono due alte maree, i ritmi vengono definiti « ritmi bimodali di giorno lunare », in contrasto coi ritmi unimodali di giorno solare degli organismi sincronizzati sul giorno solare di 24 ore. Gli orologi biologici regolati sia sul giorno lunare, sia su quello solare, hanno un'importanza notevole per la sopravvivenza, in quanto avvertono anzitempo dei cambiamenti regolari di taluni aspetti periodici dell'ambiente, come il crepuscolo o il ritorno dell'alta marea: poiché l'orologio continua a funzionare anche nelle condizioni immutabili del laboratorio, si deduce che i ritmi biologici persistano per un considerevole periodo di tempo.

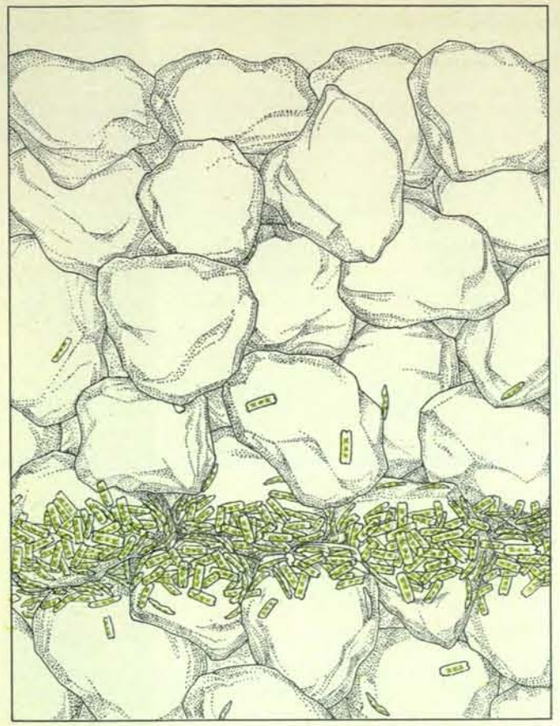
L'uca, o granchio violinista, un crostaceo adattatosi evolutivamente alle piattaforme fangose e sabbiose delle coste nordamericane, emerge dalla tana durante la bassa marea: lo si vede procedere con la sua caratteristica andatura laterale attraverso le spiagge, in cerca di detriti. I maschi ingaggiano finti combattimenti tra loro e cercano di attirare le femmine nella tana da scapoli con goffi cenni dell'enorme chela: quando arriva l'ondata di marea tutti i granchi si ritirano nella tana, dove rimangono al sicuro fino al termine dell'alta marea.

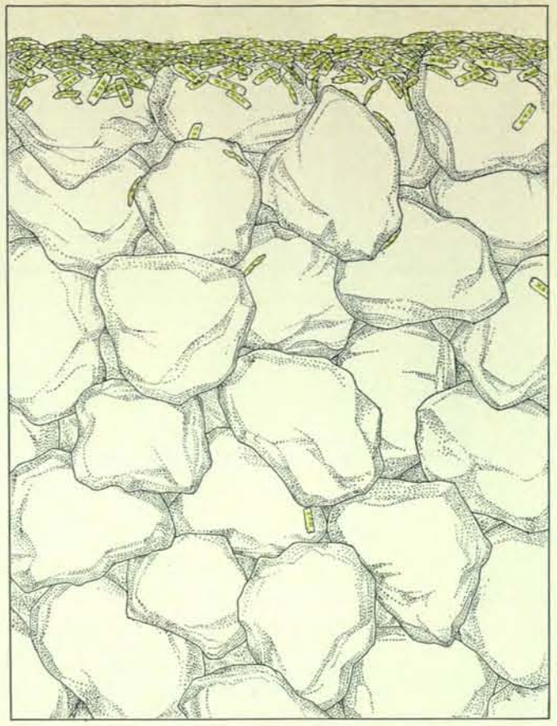
In laboratorio è molto semplice registrare il comportamento locomotore di questo granchio. Si usano come contenitori semplici scatole di plastica, come quelle in cui si pongono le esche per pescare, facilmente reperibili presso i negozi sportivi, e in ciascuna di esse si pone un granchio. La scatola, bilanciata su un fulcro a coltello, si inclina non appena il crostaceo si muove da un'estremità all'altra, provocando la chiusura di un microinterruttore, il quale, a sua volta, produce uno spostamento d'una penna che scrive su una striscia di carta. Questo registratore di movimenti viene posto in un termostato, cosicché il granchio può sbizzarrirsi negli andirivieni per giorni e giorni. In questo ambiente monotono l'orologio dell'animale continua a funzionare e a registrare quasi fedelmente (ma non del tutto) i medesimi movimenti che si verificano in natura. La differenza è lieve, ma significativa: in laboratorio il periodo del ritmo bimodale di giorno lunare è leggermente più lungo o più corto dell'analogo periodo dell'ambiente naturale. Questa variazione di periodicità si riscontra ogni volta che un organismo viene posto in un ambiente dove non si verificano cambiamenti.

Poiché i ritmi di marea seguono il giorno lunare, prendono il nome di « circalunadiani » (dal latino *circa lunarem diem*: quasi uguali a un giorno lunare).

I ritmi in alcune specie di granchi violinisti continuano inalterati talvol-

La spiaggia di Cape Cod è popolata da colonie della diatomea *Hantzschia virgata*. Questa alga durante la bassa marea vive nella sabbia umida tra la linea d'alta marea e il bordo dell'acqua. La fotografia in alto mostra una parte della spiaggia appena liberata dal riflusso della marea. Nella fotografia in basso, scattata alcuni minuti più tardi, appaiono chiazze bruno-dorate, formate dalle diatomee migrate in superficie. Quando si pone in un termostato la sabbia contenente queste diatomee, con luce e temperatura costanti, gli organismi continuano a migrare verso la superficie in sincronia coi periodi delle maree diurne che si verificano sulla spiaggia di origine. L'oggetto nelle fotografie serve da riferimento per poter confrontare le differenze di colore.





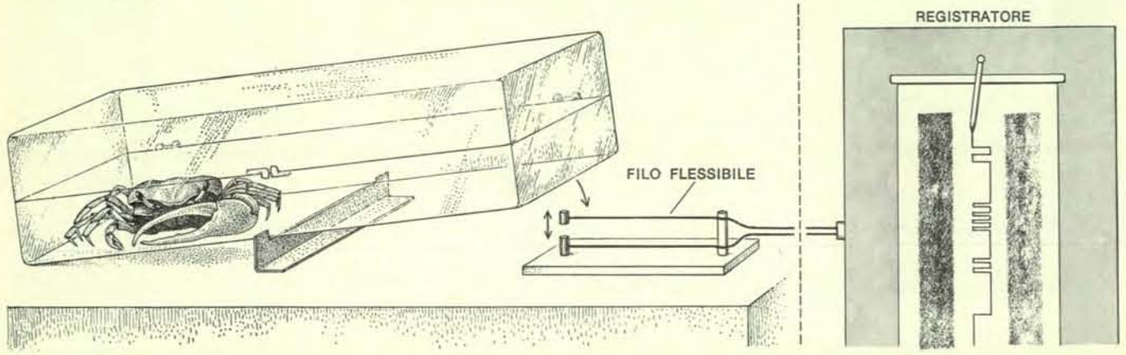
Il comportamento migratorio di *Hantzschia virgata* è schematizzato in queste due sezioni verticali di terreno sabbioso. Le diatomee di questa specie normalmente risiedono circa un millimetro sotto la superficie della sabbia (*a sinistra*). Ogni diatomea possiede due cloroplasti a forma di X, in cui avviene la fotosintesi. Durante le ore diurne gli organismi sono spinti verso la superficie dal muco espulso attraverso i pori terminali delle pareti cellulari, allungate e di consistenza silicea (*a destra*). Le diatomee rimangono esposte alla luce del sole fino al momento in cui la sabbia viene invasa dall'ondata di ritorno della marea.

ta per circa cinque settimane in laboratorio, ma più spesso si smorzano abbastanza rapidamente. I granchi debbono essere esposti di tanto in tanto a immersioni periodiche in acqua di mare, se si vuole che il ritmo mareale perduri a lungo: anche in natura, quando piccole popolazioni di granchi violinisti si stabiliscono lungo i margini di pozze non soggette a variazioni di marea, perdono il ritmo mareale e mantengono unicamente il ritmo legato al giorno solare. Quando però questi crostacei vengono riportati su una spiaggia soggetta alla marea, riescono a ristabilire velocemente il ritmo legato al giorno lunare, che poi persiste abbastanza a lungo anche quando gli animali vengono portati in località lontane.

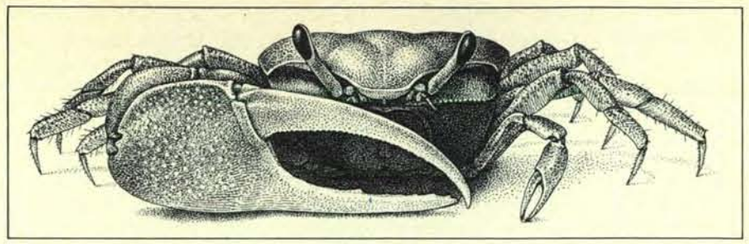
Sulle medesime piattaforme di marea vivono, a fianco del granchio violinista, due altre specie, il granchio verde e il granchio della penultima ora: entrambi questi crostacei mostrano ritmi mareali, ma differiscono dal granchio violinista in quanto presentano un'attività sincronizzata col periodo d'alta marea. I ritmi delle due specie persistono nelle condizioni costanti del laboratorio per una settimana, poi svaniscono. Nel caso del granchio verde si è potuto dimostrare che questi animali, pur avendo perso il ritmo nel laboratorio, possono recuperarlo senza essere sottoposti alle maree: basta raffreddare l'ambiente a una temperatura di 4 °C per sei ore.

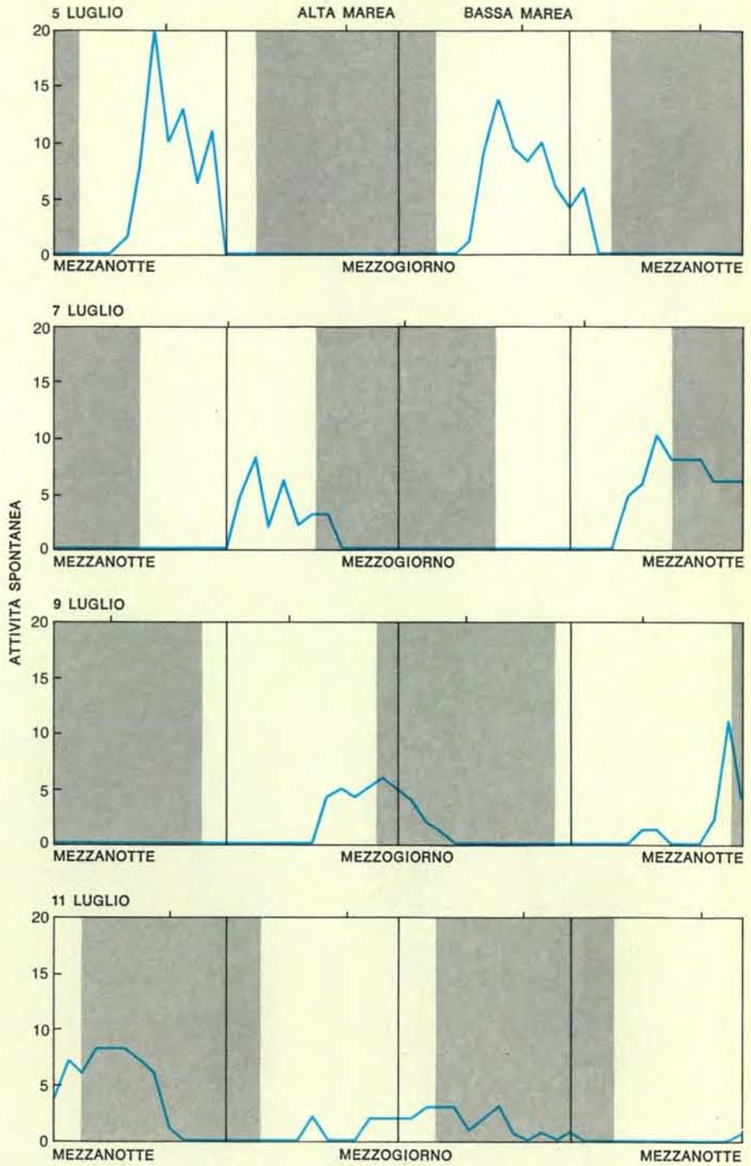
Questa tecnica è stata utilizzata anche per dimostrare che i ritmi mareali non vengono appresi dall'animale durante il processo di crescita, ma sono innati. Barbara Williams, compiendo ricerche con Ernst Naylor all'University College di Swansea, riuscì con pazienza e abilità a far nascere e allevare granchi verdi in laboratorio, partendo dalle uova, attraverso i numerosi stadi larvali, fino a giungere alla fase adulta. Durante l'intero processo di maturazione i granchi furono esposti solo ai cambiamenti periodici giorno-notte del laboratorio. Quando i granchi raggiunsero dimensioni sufficienti per essere utilizzati nei nostri esperimenti, si scoprì che l'attività locomotoria era limitata alle ore diurne. La Williams praticò a questo punto ai granchi un « trattamento a freddo » per 15 ore e registrò in seguito il comportamento locomotorio: apparve allora d'un tratto una distinta componente mareale. Poiché un periodo singolo di 15 ore di freddo non avrebbe potuto fornire ai granchi alcuna informazione sul ciclo delle maree, di durata pari a 12,4 ore, è ragionevole concludere che l'orologio che misura la frequenza mareale è innato e che per essere attivato ha bisogno solo di uno stimolo ambientale opportuno.

In condizioni naturali sia il granchio della penultima ora, sia quello verde, mostrano durante le loro attività locomotorie un evidente ritmo di giorno solare oltre al ritmo di giorno lunare. Nel granchio della penultima ora il ritmo solare appare come un largo picco d'attività che perdura per tutta la notte. Nel granchio verde il ritmo solare è rappresentato non come un unico picco, ma come una diminuzione nell'attività quando la marea diurna è massima. La combinazione dei ritmi solari e lunari, piuttosto frequente negli organismi che vivono nella zona di marea, solleva la questione se tali organismi possiedono un orologio solare per una componente del ritmo e un orologio lunare separato per l'altra componente, oppure se esiste un unico orologio che faccia funzionare entrambi i ritmi: un meccanismo a orologio singolo potrebbe essere paragonato a quel tipo d'orologio da polso usato da alcuni pescatori, che presenta contempo-



L'attività giornaliera di un granchio violinista viene registrata con un semplice strumento costituito da una scatola di plastica in equilibrio sopra un fulcro a coltello: quando il granchio sposta una estremità della scatola, questa s'inclina da quella parte e chiude un interruttore collegato a un pennino. È così possibile conoscere il numero degli andirivieni del granchio.





Il granchio violinista (*in alto*) vive sulle piattaforme fangose e sabbiose soggette alla marea: durante l'alta marea rimane nella sua tana, mentre durante la bassa marea esce in cerca di cibo. Quando uno di questi granchi viene tolto dal suo ambiente e posto in un termostato con luce e temperatura costanti, all'inizio i suoi periodi di massima attività corrispondono ai tempi di bassa marea della località di origine (*curva in alto*): poi il periodo del ritmo comincia ad allungarsi e verso la fine della settimana non è più in sincronia con i tempi di marea della spiaggia di origine (*curva in basso*).

raneamente sul quadrante l'indicazione dell'ora solare e dell'ora di marea.

Oltre all'attività locomotoria, nei granchi esistono anche altri processi regolati da un orologio biologico. I ritmi di variazione di colore del granchio violinista, del granchio verde e di quello della penultima ora sono stati studiati da Frank A. Brown jr., Marguerite Webb e Milton Fingerman del Marine Biological Laboratory e da B.L. Powell del Trinity College di Dublino. Nell'ipoderma di questi granchi sono presenti dei cromatofori stellati che contengono granuli d'un pigmento scuro: quando i granuli di pigmento sono strettamente aggregati nel centro delle cellule, la colorazione dei granchi è chiara, quando i granuli sono uniformemente dispersi in tutti i prolungamenti delle cellule, l'effetto di colore che risulta è scuro. Tutte e tre le specie impallidiscono di notte e diventano scure durante le ore diurne, anche quando sono poste nelle condizioni costanti di un laboratorio. Si è trovato che la variazione di colore del granchio violinista possiede anche una componente regolata dalla marea che produce un oscuramento addizionale al momento della bassa marea.

Gli occhi dei granchi sono posti su peduncoli mobili, sedi di un'unità neuroendocrina chiamata complesso organo X - ghiandola del seno, che secerne un ormone in grado di far disperdere i granuli di pigmento all'interno dei cromatofori. Powell ha dimostrato che l'asportazione dei peduncoli oculari del granchio verde (e quindi anche del complesso organo X - ghiandola del seno) elimina il ritmo di cambiamento di colore del granchio. Inoltre lo stesso ricercatore ha potuto dimostrare che il ritmo si ripristina in un granchio senza peduncoli, innestandogli le ghiandole del peduncolo d'un altro granchio. Questi risultati dimostrano che nel granchio verde i peduncoli oculari sono la sede dell'orologio che controlla il ritmo di mutamento di colore: nel granchio della penultima ora e nel granchio violinista l'asportazione dei peduncoli oculari riduce invece unicamente l'ampiezza dei ritmi di cambiamento di colore.

Il sistema neuroendocrino dei peduncoli oculari esercita anche un controllo sul ritmo di attività locomotoria del granchio verde. Negli esperimenti condotti da Naylor e Williams, quando i peduncoli oculari dei granchi verdi venivano asportati, tutta l'attività locomotoria cessava e ritornava solo gradualmente nei successivi sei giorni. Poiché nei granchi verdi il ritmo di marea svanisce normalmente dopo circa una settimana nelle condizioni costanti del laboratorio, la mancanza di ritmo nei granchi senza peduncoli quando riprendevano la loro attività non era affatto strana. Tuttavia i tentativi per restaurare il ritmo, mediante immersione dei granchi in acqua fredda, non ebbero alcun effetto. D'altro canto i granchi a cui erano state asportate solo le retine, e non i peduncoli, ritornavano al ritmo normale dopo essere stati immersi in acqua fredda.

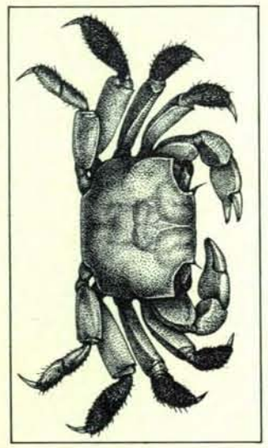
In successivi esperimenti Naylor e Williams trovarono che non era necessario sottoporre tutto il corpo del granchio al tuffo nell'acqua fredda. Se si pone un granchio aritmico in acqua con una temperatura di 15 °C, ma legato in modo da fare sporgere i peduncoli oculari sopra la superficie, il suo ritmo mareale può essere ripristinato semplicemente facendo gocciolare acqua marina ghiacciata sopra i peduncoli per un breve periodo. Infine, Naylor e Williams ricavarono estratti di peduncoli oculari di granchi verdi nella fase quiescente del ritmo di attività locomotoria e iniettarono gli estratti in granchi attivi a cui erano stati estirpati i peduncoli. L'iniezione provocò una riduzione significativa del livello di attività di questi granchi, dimostrando che esiste una sostanza inibitrice liberata periodicamente dalle ghiandole del peduncolo.

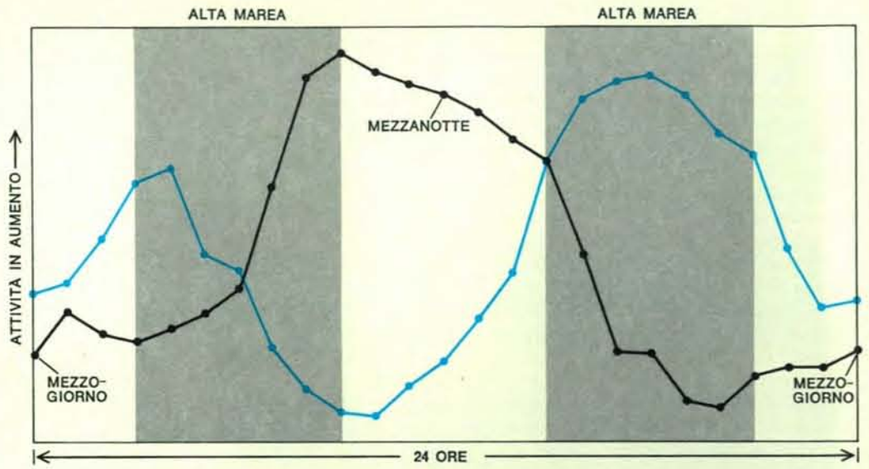
Ho effettuato esperimenti simili sul granchio della penultima ora. Quando i peduncoli oculari di questo granchio venivano asportati, tutta l'attività locomotoria si fermava. Ricavai estratti di peduncoli oculari di granchi ritmici sia durante la fase attiva, sia durante la fase quiescente dell'attività locomotoria e li iniettai in varie concentrazioni in granchi che erano diventati aritmici a causa di una permanenza prolungata in condizioni costanti. In seguito a questo esperimento non si poterono osservare alterazioni consistenti nei livelli di attività dei granchi recettori.

Poiché le ghiandole neuroendocrine dei peduncoli oculari del granchio della penultima ora non determinano il ritmo dell'attività locomotoria di questa specie, effettuai un esperimento per scoprire se vi fosse un messaggero chimico proveniente da qualche altro punto del corpo: congiunsi due granchi, uno dei quali fortemente ritmico, l'altro aritmico, ritagliando due piccole aperture nell'esoscheletro dorsale dei due animali e saldando queste ferite con la ceralacca. Nei granchi la maggior parte del sangue non è contenuta in vasi sanguigni, ma fluisce liberamente negli spazi tra gli organi: quindi, quando due granchi restano congiunti nel modo che ho descritto sopra, il sangue dell'uno si mescola liberamente con quello dell'altro.

Tenni presente anche una particolarità anatomica dei granchi, il processo detto autotomia. Quando un granchio viene aggredito, l'attaccante di solito gli afferra una delle 10 zampe: la vittima si difende eliminando la zampa e scappando prima che il predatore gliene afferri un'altra. La zampa si separa dal corpo in un punto di rottura predeterminato: la perdita eccessiva di sangue dal moncone viene evitata mediante un meccanismo di blocco automatico e la zampa sacrificata si rigenera durante le mute successive.

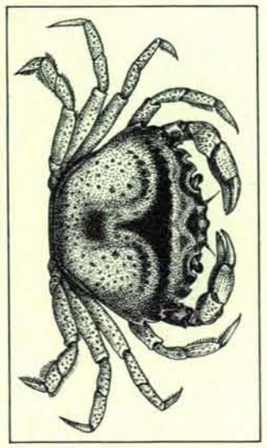
Sfruttando questo meccanismo di amputazione automatica, poco prima di congiungere i due granchi, indussi il granchio ritmico a disfarsi di tutte le sue zampe. Perciò dei due granchi così congiunti uno, aritmico, stava in basso e camminava, mentre l'altro, ritmico ma privo delle zampe, si faceva trasportare passivamente, rimanendo, per forza di cose, supino (*si veda l'illustrazione a pagina 70*). Qualsiasi attività locomotoria registrata da questo momento sarebbe stata l'attività dell'individuo fornito di zampe e avrebbe avuto il significato di attività ritmica indotta da una sostanza ignota, presente nel sangue del granchio amputato. Nei 47 esperimenti condotti su coppie congiunte, non potei registrare alcun ritmo. D'altra parte, quando due granchi ritmici venivano saldati nello

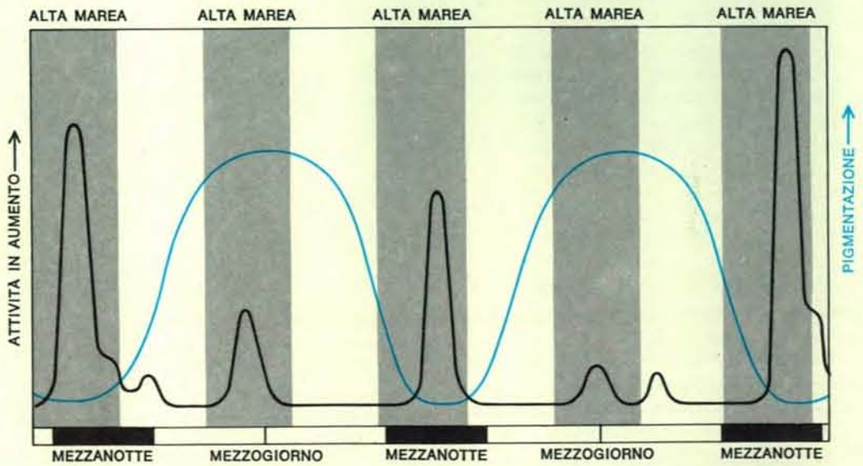




Il granchio della penultima ora vive sulle piattaforme di marea in una zona molto vicina a quella del granchio violinista: il suo nome deriva dal fatto che l'attività di un individuo appena catturato raggiunge il culmine un'ora prima della mezzanotte.

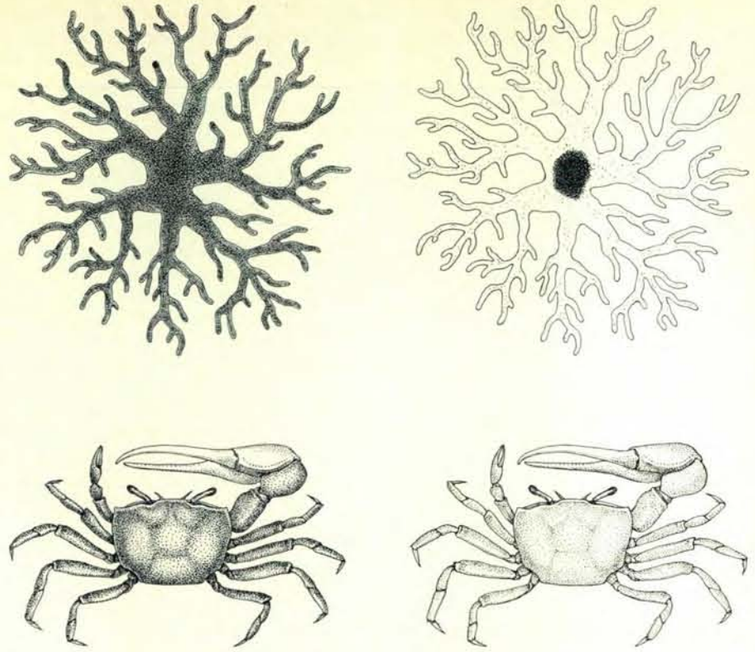
La curva in nero mostra l'attività media regolata sul giorno solare, di 30 granchi per un periodo di un mese. Lo stesso crostaceo dimostra anche un ritmo di attività regolato sul giorno lunare (*curva in colore*) che corrisponde ai periodi d'alta marea.





Il granchio verde, che vive pure sulle piattaforme soggette a maree, possiede un modello di attività locomotoria corrispondente ai periodi d'alta marea. L'attività diminuisce nettamente quando l'alta marea si verifica durante le ore diurne. Questo comportamento non viene perduto quando il granchio viene portato in laboratorio in condizioni costanti (*curva in nero*).

Il granchio verde mostra inoltre un ritmo, regolato dal giorno solare, di variazione del proprio colore: impallidisce durante la notte e diviene più scuro durante le ore diurne. Il ritmo di variazione cromatica (*curva in colore*), che persiste quando il granchio è mantenuto in condizioni costanti, si pensa sia controllato da un sistema neuroendocrino localizzato nei peduncoli oculari.



La variazione di colore del granchio violinista è il risultato di un'aggregazione o di una dispersione di granuli di pigmento in alcune cellule dell'ipoderma, dette cromatofori (*in alto, notevolmente ingranditi*): in alcuni cromatofori vi sono solo granuli neri, in altri anche granuli bianchi o arancioni. Questa specie di granchio diventa più chiara durante la notte mentre si scurisce in presenza di luce diurna. La componente mareale di questo ritmo fa sì che vi sia un oscuramento addizionale nel periodo di bassa marea.

stesso modo in esperimenti di controllo, si poteva osservare sempre un ritmo: quindi il procedimento di fusione non era responsabile dell'assenza di ritmicità.

È chiaro quindi che, mentre il sistema endocrino è coinvolto nei ritmi di variazioni di colore e di locomozione nel granchio verde, non è necessariamente coinvolto in ritmi di questo genere in altri granchi. Non è neppure sicuro che un meccanismo neurale stia alla base di un qualsiasi ritmo biologico, poiché un livello di organizzazione unicellulare come quello della diatomea *Hantzschia* è sufficiente a esprimere tutte le note proprietà dei ritmi regolati da un orologio biologico.

L'attitudine al ritmo non viene appresa e neppure rimane impressa su un organismo per effetto dell'ambiente circostante, ma è l'espressione d'un fattore potenziale genetico. L'ereditarietà determina anche se il granchio deve essere attivo durante l'alta o la bassa marea: questo non significa per altro che l'ambiente non abbia un ruolo importante nella chiara manifestazione di un ritmo. L'orario delle maree su un particolare tratto di costa determina l'adattamento del ritmo. La relazione tra un orologio biologico e l'ambiente è simile a quella che si stabilisce tra un orologio a pendolo e il suo proprietario. Il ritmo con cui cammina un orologio a pendolo è determinato dai suoi meccanismi e dal tipo di pendolo, ma il proprietario può spostare in ogni momento le lancette con le dita per riportarle sull'ora esatta. Così un granchio verde si adatterà ben presto a un massimo d'attività durante l'alta marea e un granchio violinista durante la bassa marea, anche dopo che questi due animali siano stati trasportati su una lontana spiaggia d'un differente oceano.

Nella sabbia che forma le dune delle spiagge della California meridionale è diffuso un piccolo crostaceo saltatore del genere *Excirolana*. Al culmine dell'alta marea, quando le acque inondano l'habitat di questo minuscolo isopodo, l'animaletto emerge dalla sabbia per nuotare e andare in cerca di cibo tra i frangenti. Due o tre ore più tardi, quando la marea si ritira, il crostaceo si rintana nella sabbia e aspetta la successiva ondata di marea. James T. Enright della Scripps Institution of Oceanography trovò che tali crostacei, tenuti in un vaso d'acqua marina in condizioni costanti, nuotavano attivamente durante i periodi corrispondenti al massimo d'alta marea e rimanevano in riposo sul fondo del recipiente per tutto il resto del tempo.

Nella California meridionale l'andamento delle maree varia notevolmente secondo le fasi lunari. In un determinato mese il tipo di marea passa da un massimo per ogni giorno lunare a due per giorno lunare. Inoltre, durante i periodi di transizione da una marea al giorno a due e da due a una, l'altezza dei massimi di marea consecutivi varia. L.A. Klapow, che allora lavorava all'Università di California a San Diego, dimostrò che l'andamento delle maree nel periodo in cui i piccoli crostacei erano stati raccolti veniva riflesso dall'intensità del ritmo di attività degli stessi animali portati in laboratorio (*si veda l'illustrazione nella pagina a fronte*). In esperimenti separati Klapow ed Enright dimostrarono anche che le onde battenti e le acque turbinose determinano il modello di ritmo dell'isopodo.

Sembra perciò che gli abitatori delle spiagge esposte al mare aperto abbiano modelli di attività regolati dall'azione dei frangenti. Gli organismi che vivono in baie protette nella zona di marea non sono invece esposti a frangenti e perciò dobbiamo cercare altri fattori che ne determinano il ritmo; le possibilità sono numerose: l'inondazione periodica e le variazioni regolari di temperatura, la pressione idrostatica, la composizione chimica delle acque o la disponibilità d'ossigeno. Di tutti questi possibili fattori, solo due si sono dimostrati veramente importanti: come esempio mi basterà citare un altro studio sul granchio verde, eseguito dall'infaticabile gruppo di ricercatori condotti da Naylor e Williams.

Una delle scoperte più sorprendenti è stata quella che la principale caratteristica della marea, la periodica inondazione della linea di costa, non era di per se stessa un fattore importante nel sincronizzare il ritmo di attività locomotoria con le maree. Questo fatto fu dimostrato portando dei granchi in laboratorio e sottoponendoli per cinque giorni a 6,2 ore d'immersione in acqua marina, seguiti da 6,2 ore di esposizione all'aria. L'immersione in acqua marina fu calcolata in modo da corrispondere alla bassa marea nella spiaggia natia dei granchi, capovolgendo così l'orario di marea di questi crostacei. La temperatura dell'acqua e dell'aria fu mantenuta costante a 19 °C. Dopo questo trattamento i granchi furono posti in un registratore d'attività, dove i movimenti furono misurati durante i successivi tre giorni mantenendo costante la temperatura. Il trattamento non riuscì a riportare in fase il ritmo dei granchi. Il procedimento venne ripetuto, ma stavolta la temperatura dell'aria fu mantenuta a un livello di 11 gradi più elevato di quello dell'acqua. Cinque giorni di questo trattamento rimisero in fase il ritmo dei granchi, e la variazione rimase inalterata anche in condizioni costanti.

In una versione finale del medesimo esperimento fu tralasciata l'immersione in acqua di mare. I granchi furono esposti all'aria a 13 °C per 6,2 ore e poi all'aria a 24 °C per lo stesso periodo. Ne risultò una sincronizzazione completa e persistente. Recentemente ho provocato gli stessi cambiamenti di comportamento nei granchi violinisti e in quelli della penultima ora, e ne ho dedotto che è l'abbassamento di temperatura portato dal flusso di marea che ha la maggiore importanza nello stabilire la fase del ritmo di attività dei granchi.

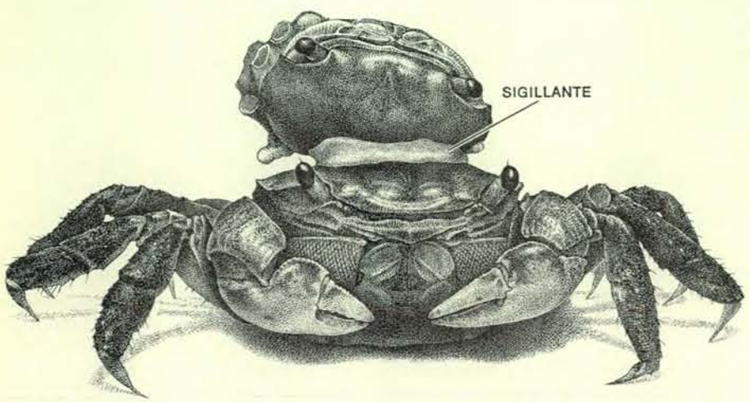
L'altra forza ambientale che sincronizza i ritmi alle maree locali è la pressione idrostatica. In un esperimento alcuni granchi aritmici furono sottoposti per cinque giorni a un ciclo di alta pressione (1,6 atmosfere) per 6,2 ore, seguito da un periodo di 6,2 ore di pressione normale. I granchi reagirono con un aumento di attività durante i periodi d'alta pressione: questa periodicità continuò quando i granchi furono mantenuti in condizioni costanti.

Per ora non abbiamo una conoscenza di prima mano di come l'orologio vivente funzioni realmente: mentre cerchiamo attivamente questo meccanismo misterioso, abbiamo se non altro chiarito molte delle sue proprietà.

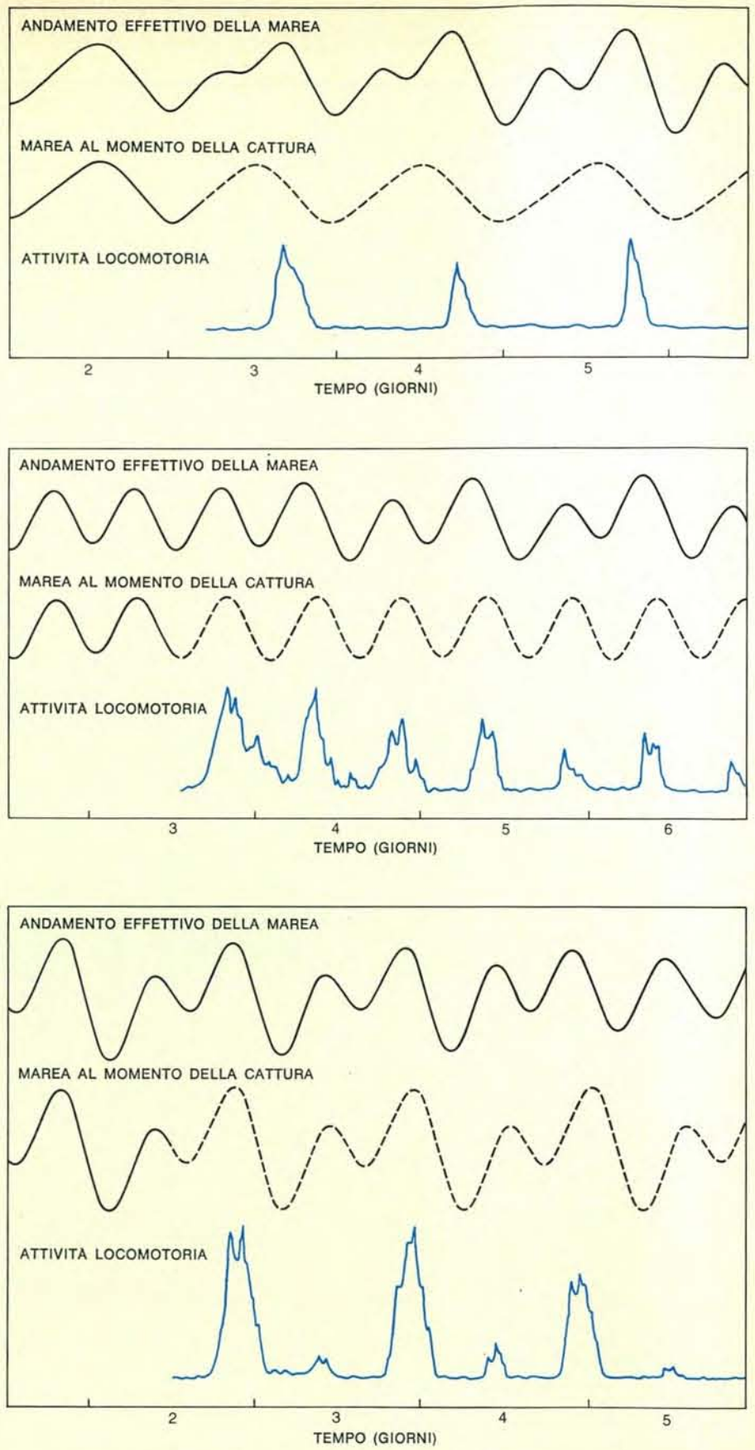
Quando furono scoperti per la prima volta i ritmi dei processi biologici come il metabolismo ossidativo, la fotosintesi e altri processi simili, e venne notato che questi ritmi persistevano in assenza di stimoli esterni, l'orologio regolatore fu interpretato semplicemente come un passaggio oscillante nella catena di reazioni chimiche implicate nei processi. Di conseguenza i primi tentativi di localizzare l'orologio si occuparono dell'analisi delle reazioni delle catene, della speranza di identificare il punto d'oscillazione. Tale componente ritmica non fu però scoperta e gli esperimenti successivi mostrarono che essa probabilmente non esiste: in effetti ora si sa che l'orologio è completamente distinto dal processo che viene reso ritmico.

Una delle numerose osservazioni che portano a questa conclusione fu eseguita sul granchio verde. Quando la temperatura corporea del granchio fu abbassata a 10 °C, tutta l'attività locomotoria si arrestò per la durata del raffreddamento. Quando la temperatura corporea fu fatta ritornare al livello normale, l'attività riprese e il ritmo di movimento si trovò esattamente in fase con quello dei granchi di controllo che non erano stati raffreddati. È chiaro che l'orologio del granchio aveva continuato a funzionare con precisione anche quando il ritmo non era stato esplicato. Questo dato dimostra che l'orologio e i processi ritmici prodotti sono separati e debbono essere congiunti in modo da poter essere staccati e poi riagganciati.

Il distacco della coppia orologio-processo indotto è forse anche la causa della perdita finale di ritmicità che si



In questa figura due granchi della penultima ora appaiono uniti in modo che il loro sangue si può mescolare liberamente: nell'esoscheletro dorsale di ciascun granchio vengono effettuate alcune piccole incisioni, che poi vengono saldate con ceralacca. Il granchio in alto si è liberato di tutte le sue zampe per mezzo dell'autotomia, un processo mediante il quale un granchio stacca automaticamente l'arto afferrato da un predatore. Prima dell'unione l'attività locomotoria del granchio così amputato era sincrona con le maree, mentre l'attività del granchio illeso, era aritmica. Negli esperimenti condotti su queste coppie di granchi non si è mai rilevata un'attività ritmica: si conclude perciò che il sangue non contiene alcun messaggero chimico che induce tale comportamento.



Il ritmo di attività del genere *Excirolana*, un piccolo crostaceo isopodo detto pulce della sabbia che vive sulle spiagge sabbiose della California meridionale, si è adattato alla peculiarità delle maree nella regione. Le maree si alternano ogni mese tra un massimo per giorno lunare e due massimi. Quando i piccoli isopodi vengono mantenuti nelle condizioni costanti del laboratorio, il loro tipo di attività (*curve in colore*) tende a imitare l'andamento delle maree a cui sono stati esposti per l'ultima volta (*curve tratteggiate*) anziché l'andamento effettivo delle maree (*curve continue in nero*).

nota negli animali mantenuti a lungo nelle condizioni costanti del laboratorio. Sembra che, quando gli animali della zona di marea sono portati lontano dal loro ambiente, non vi è più alcuna forza che li mantenga nel ritmo mareale e quindi i processi vitali, svincolati dall'orologio, divengono aritmici.

La validità della nozione di accoppiamento tra questo orologio e i processi vitali è ulteriormente messa in rilievo dal fatto che i ritmi, scomparsi negli animali in laboratorio o in un habitat naturale non soggetto a marea, si possono ripristinare per mezzo di uno stimolo semplice e di breve durata, per esempio il raffreddamento. Poiché il trattamento non fornisce alcuna informazione circa gli intervalli di marea, l'interpretazione più semplice è che lo stimolo riagganci la coppia formata dall'orologio, che nel frattempo ha continuato a funzionare, e i processi che regolano l'attività locomotoria: tale attività quindi ridiventa ritmica.

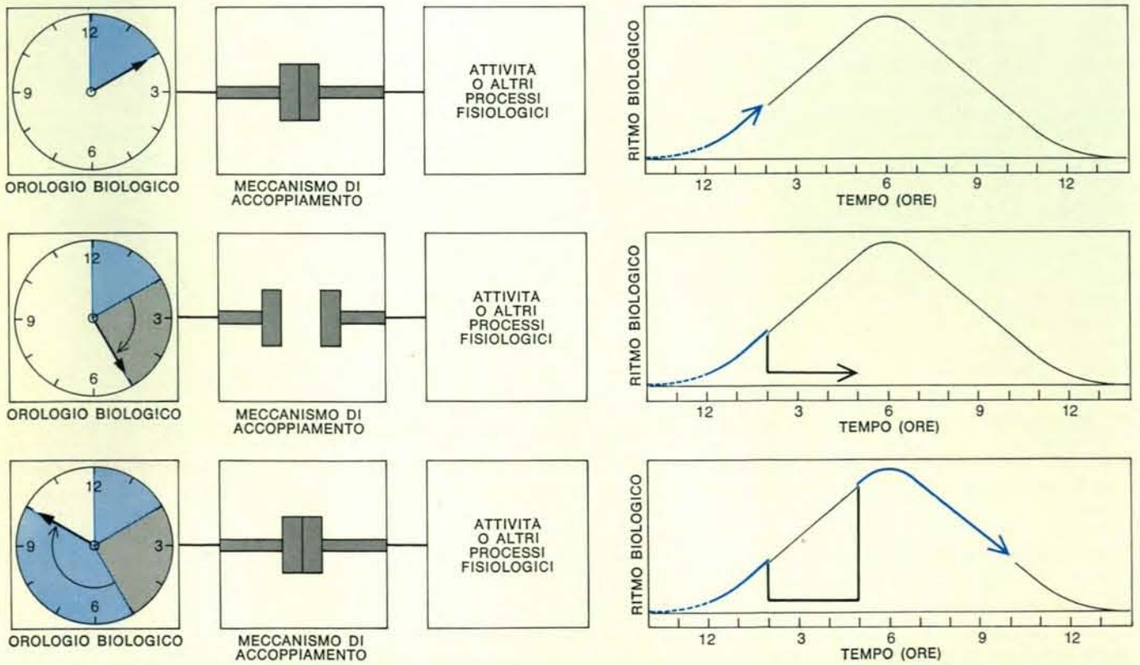
Come abbiamo visto, il ritmo di migrazione verticale della diatomea *Hantzschia* dimostra che a un orologio biologico è sufficiente il livello di organizzazione caratteristico di un'unica cellula per potersi esprimere. Esistono altri due organismi unicellulari che forniscono esempi ancor più evidenti. Il dinoflagellato marino *Gonyaulax* dimostra di possedere simultaneamente ritmi diversi in quattro processi: fotosintesi, luminescenza (emissione di luce durante la notte), irritabilità e divisione cellulare. Nell'alga verde unicellulare *Acetabularia* sono stati scoperti ben cinque ritmi diversi, che persistono anche quando il nucleo della cellula viene asportato con un metodo microchirurgico. Esistono prove che nelle piante e negli animali pluricellulari l'orologio debba essere ricercato nelle singole cellule: quando gli organismi sono suddivisi nelle loro parti e alcuni pezzi vengono mantenuti in vita in coltura di tessuto, le cellule continuano il ritmo originario. Il luogo più plausibile in cui si deve ricercare questo orologio vivente è quindi all'interno della singola cellula, dove ci si può aspettare di scoprirlo sotto forma di una qualsiasi entità fisiologica. Tuttavia, nonostante ricerche approfondite, non sono stati localizzati né l'orologio, né alcuna delle sue componenti. La ricerca comunque ha rivelato due aspetti insoliti di questo meccanismo: la velocità con cui funziona è quasi completamente insensibile alla temperatura e non è neppure influenzata da tutta una gamma di agenti chimici potenzialmente demolitori.

In generale aumentando la temperatura aumenta la velocità con cui una reazione chimica procede: ci si aspetterebbe quindi che l'orologio vivente, con tutto il suo apparato chimico, acceleri la propria velocità in modo analogo. Per sottoporre a esperimento queste ipotesi di lavoro, abbiamo sottoposto gruppi di granchi a temperature gradualmente superiori in laboratorio e abbiamo osservato i loro ritmi in periodi di parecchi giorni. Poiché il periodo d'un ritmo biologico espresso si considera corrispondente alla frequenza portante dell'orologio, un cambiamento del ritmo indotto dalla temperatura si può considerare corrispondente a un mutamento nella frequenza dell'orologio. Il risultato ottenuto negli esperimenti indicati è l'assenza completa di variazioni nel periodo: se si osserva una variazione, si tratta soltanto di una frazione minima delle variazioni normali nei sistemi chimici.

Quasi nello stesso modo si sono dimostrati infruttuosi i tentativi aventi lo scopo di demolire i ritmi degli organismi con sostanze chimiche, come inibitori della sintesi proteica, stimolanti, inibitori metabolici e sostanze narcotizzanti. Delle centinaia di sostanze utilizzate negli esperimenti, solo quattro alterano il periodo del ritmo: l'ossido di deuterio, l'alcool etilico, la valinomicina e gli ioni di litio. Data la grande varietà e il numero di sostanze vagliate, appare chiaro che gli orologi biologici, a differenza degli altri sistemi autoregolatori degli organismi, sono virtualmente immuni a tutte le manipolazioni chimiche.

Da un punto di vista pratico questa insensibilità avrebbe potuto essere prevista. È certo che uno dei più importanti attributi di qualsiasi orologio, vivente o artificiale, è la precisione: un orologio la cui velocità venisse alterata da variazioni di temperatura o da sostanze chimiche presenti nell'ambiente non avrebbe i requisiti opportuni. Anzi, se l'orologio reagisse a ogni cambiamento della temperatura ambientale, non sarebbe a rigor di termini un orologio, ma piuttosto un termometro che segnala le temperature ambientali mediante la velocità con cui corre.

La precisione degli orologi biologici è ancor più stupefacente quando ci si rende conto del fatto che tale esattezza dev'essere mantenuta durante la divisione cellulare, quando presumibilmente non solo la cellula ma anche l'orologio viene riprodotto. La regolarità con cui viene compiuta questa riproduzione è stata dimostrata in una ricerca sul paramecio, un protozoo unicellulare, eseguita da Audrey Barnett dell'Università del Maryland. Nel ceppo di parameci allevati dalla studiosa, il sesso di ciascun animale varia ogni giorno, nel periodo dell'accoppiamento, da un individuo all'altro. La Barnett sottopose a oscurità continua otto parameci i cui orologi erano stati regolati sui cicli normali giorno-notte. Le cellule della popolazione si dividevano 2,2 volte al giorno, e alla fine dei sei giorni avevano prodotto più di 121 000 individui. Il settimo giorno fu di nuovo esaminato il comportamento di inversione sessuale dell'intera popolazione, e si trovò che era ritmico. La fase del ritmo era assai simile a quella delle cellule di controllo rimaste nel ciclo normale giorno-notte. Poiché solo le otto cellule originarie erano state sottoposte al ciclo giorno-notte, ciascuno degli orologi originari era stato riprodotto numerose volte con una perdita minima di precisione. Il fatto può essere anche interpretato in questo senso: ogni cellula contiene molti orologi alcuni dei quali si riproducono mentre altri rimangono ancora collegati ai processi cellulari e ne regolano il ritmo.



Si pensa che esista un meccanismo di accoppiamento, non ancora identificato, tra gli « orologi » e i processi biologici. Quando la coppia è agganciata (*a*), si ha l'espressione del ritmo biologico. Lo sganciamento della coppia (*b*) può causare la perdita del comportamento ritmico degli organismi allevati in laboratorio. Tuttavia l'orologio continua a funzionare e, quando il meccanismo viene riagganciato (*c*), il ritmo ricomincia al punto corrispondente al « tempo » determinato dall'orologio.

EDWARDS "IL VUOTO NEL MONDO"

Infatti il nome Edwards è sinonimo di tecnica del vuoto,
e chi opera in questo campo lo sa.
Il Gruppo Edwards ha una propria sede o consociata in tutti i più importanti paesi del mondo.
Per essere più vicino ai Vostri problemi.
Per rispondere prontamente alle Vostre richieste.
E per Voi, nel Gruppo Edwards, lavorano uomini specializzati.
Nel settore del vuoto, del freddo, della meccanica, della strumentazione elettronica, della liofilizzazione farmaceutica.
Ed è questa specializzazione che ci ha permesso di assumere l'attuale proporzione mondiale.
Per essere sempre proiettati alla ricerca del meglio, ci siamo suddivisi i compiti di studio e progettazione.
In Italia, gli impianti di liofilizzazione sono la nostra prima ragione di vita.
Li progettiamo e li costruiamo per Voi e per tutto il mondo.

Edwards
GRUPPO BOC

EDWARDS ALTO VUOTO S.p.A.
20090 TREZZANO SUL NAVIGLIO (MI) VIA CARPACCIO 35 - TEL. 44.53.851
TELEX 32345 EDHIVAC - TELEGR. EDVUOTO TREZZANO S/N
UFFICIO REGIONALE: 00154 ROMA - VIA CAPITAN BAVASTRO 136 - TEL. 57.61.58

Alfred Wegener e l'ipotesi della deriva dei continenti

Sessant'anni fa questo scienziato avanzava l'ipotesi che i continenti si muovono e proponeva una teoria della loro migrazione; la validità delle sue idee, tuttavia, non venne riconosciuta se non molto più tardi

di A. Hallam

Che le linee di costa dei due lati dell'oceano Atlantico combaciano l'una con l'altra come i pezzi di un rompicapo dev'essere stato notato non appena furono compilate le prime buone carte geografiche del Nuovo Mondo. Ben presto la complementarietà della forma dei due continenti provocò delle speculazioni circa la loro origine e le loro storie, e un certo numero di teorie primitive cominciò a suggerire che queste forme non fossero tali per una semplice coincidenza. Nel 1620 Francesco Bacone richiamò l'attenzione sulla somiglianza dei contorni dei due continenti, ma non giunse a suggerire che essi potevano un tempo aver costituito un'unica massa continentale. Nei secoli successivi altre teorie cercarono di spiegare questa corrispondenza, in genere come risultato di qualche supposta catastrofe, quale lo sprofondamento della mitica Atlantide. Il primo che suggerì che i continenti si erano effettivamente mossi sulla superficie della Terra fu Antonio Snider Pellegrini nel 1858, ma anch'egli attribuì l'evento a una causa soprannaturale: il diluvio universale. Oggi, naturalmente, la migrazione dei continenti costituisce una parte essenziale della teoria della struttura della Terra universalmente accettata dai geologi.

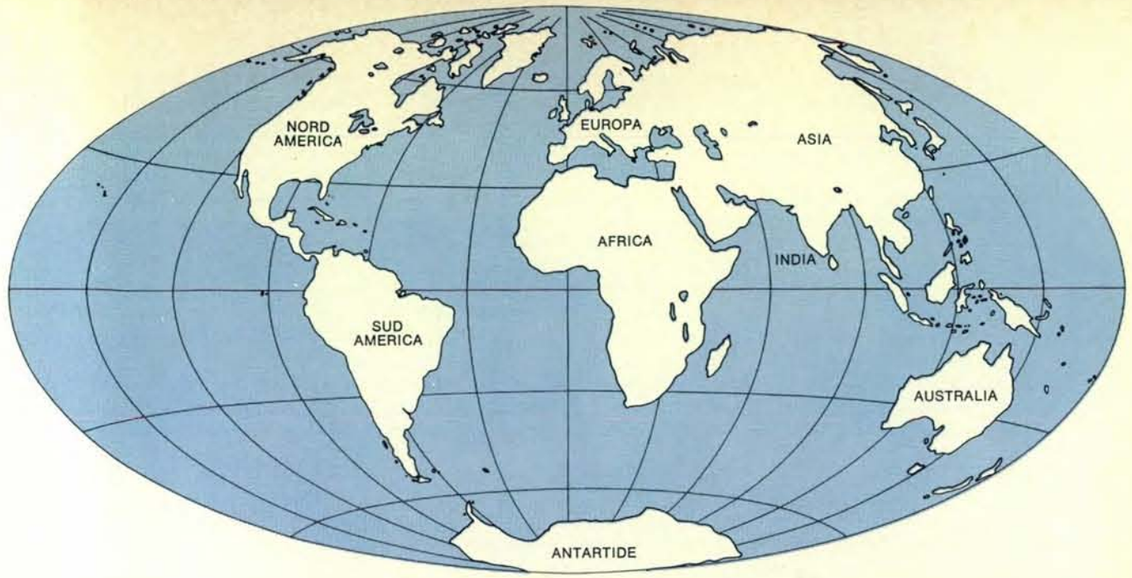
Fra le varie ipotesi che precedettero la teoria moderna della tettonica a zolle (si veda il volume *Tettonica a zolle e continenti alla deriva,* Letture da « Le Scienze », 1974), una teoria si pone in evidenza: quella formulata da Alfred Wegener nei primi anni del XX secolo. Wegener aveva accesso soltanto a una piccola parte delle informazioni disponibili oggi, eppure la sua teoria anticipa di molto alcuni aspetti fondamentali della concezione attuale della Terra: non solo il moto dei continenti, ma anche la deriva polare coi cambiamenti climatici che ne conseguono e il significato della distribuzione di antiche piante e animali. Nel suo lavoro vi sono anche parti che si sono poi dimostrate sbagliate, ma i punti fondamentali delle sue argomentazioni sono stati confermati. La sua teoria non resta semplicemente quella di un precursore dei concetti che oggi prevalgono, ma è quella del suo autentico progenitore.

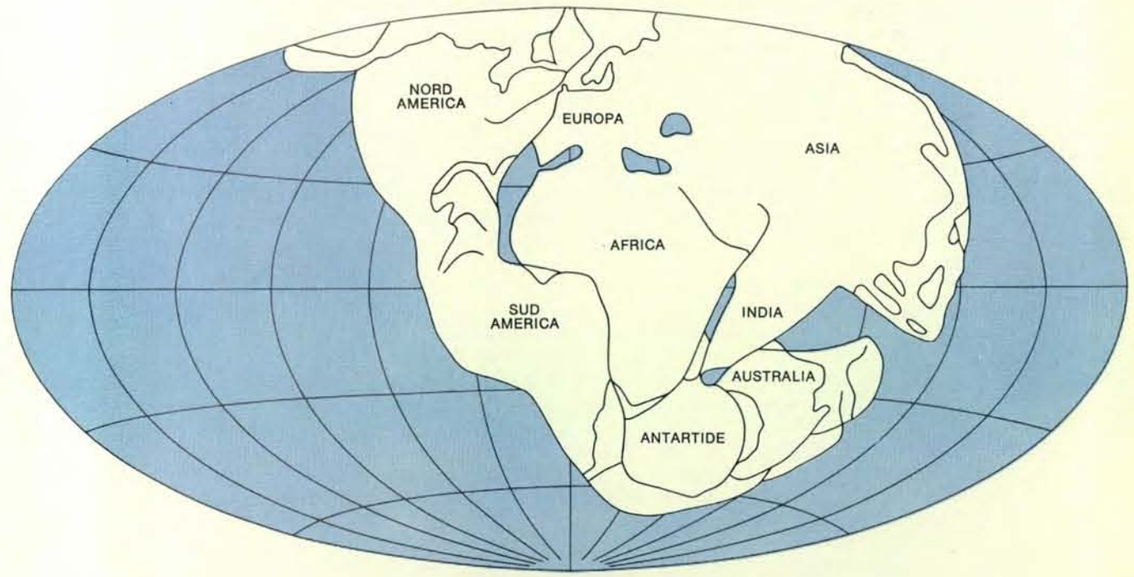
Wegener presentò per la prima volta le sue idee nell'ambiente scientifico nel 1912, ma tali idee hanno ricevuto credito generale soltanto cinquant'anni dopo. Quando infine le sue vedute sulla Terra sostituirono il modello precedente, negli anni sessanta, il cambiamento rappresentò una revisione radicale di una dottrina ben stabilita, ed esso ebbe luogo soltanto perché nuovi dati, derivati dalle scoperte in geofisica e oceanografia, lo rendevano obbligatorio. Nel frattempo la teoria di Wegener era stata quanto meno trascurata, e spesso era stata trattata con noncuranza. Nel momento peggiore i sostenitori della deriva dei continenti venivano sprezzantemente liquidati come maniaci. Per capire questa reazione bisogna esaminare sia l'opera di Wegener che l'atteggiamento dei suoi contemporanei. Ricavò Wegener le sue conclusioni da dati attendibili e le sostenne con argomenti coerenti? Oppure si limitò a formulare delle congetture e si trovò ad avere ragione per caso? Se i suoi ragionamenti erano plausibili, perché la sua opera fu avversata con tanta persistenza e determinazione? Mi proverò a rispondere a queste domande e in più cercherò di valutare che genere di uomo fosse Wegener e quale fosse il suo valore di scienziato.

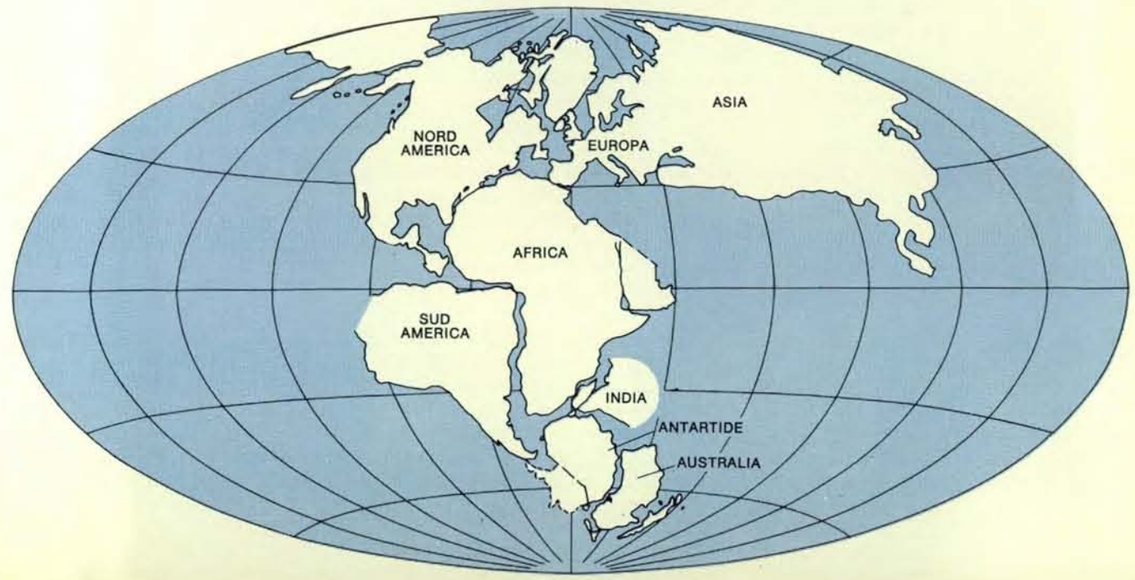
Wegener non aveva credenziali come geologo. Nato a Berlino nel 1880, figlio di un ministro evangelico, studiò alle Università di Heidelberg, Innsbruck e Berlino e ottenne il dottorato in astronomia. Sua passione predominante fu però la meteorologia.

Tale sua passione fu ricompensata quando egli fu selezionato come meteorologo per una spedizione danese nella Groenlandia nord-orientale. Al suo ritorno in Germania accettò un incarico in meteorologia all'Università di Marburg ed entro pochi anni scrisse un importante testo sulla termodinamica dell'atmosfera. Nel 1912 partecipò a una seconda spedizione in Groenlandia, col danese J.P. Koch. Questa spedizione comprendeva la più lunga traversata della calotta glaciale mai intrapresa; le osservazioni glaciologiche e meteorologiche pubblicate riempiono molti volumi.

Nel 1913 Wegener sposò Else Köppen, figlia del meteorologo Wladimir Peter Köppen. Dopo la prima guerra mondiale (in cui Wegener combatté come ufficiale) sostituì il suocero alla direzione del Dipartimento di ricerche meteorologiche dell'Osservatorio marino di Amburgo. Nel 1924 accettò una cattedra di meteorologia e geofisica all'Università di Graz in Austria, dove trovò che i suoi colleghi mostravano







Un antico supercontinente avrebbe riunito tutte le grandi masse di terra emersa del globo. L'illustrazione al centro nella pagina a fianco illustra la ricostruzione di Wegener di questo supercontinente, che egli per primo chiamò Pangea. Una versione più recente, illustrata in basso, differisce in alcuni dettagli nella posizione e nell'orientamento, ma conserva le caratteristiche principali della proposta iniziale di Wegener. Ambedue le carte si basano non sulle linee di costa dei continenti, ma sui bordi delle piattaforme continentali. Per confronto in alto è riportata una carta del mondo come appare oggi.

maggior simpatia per i suoi interessi di ricerca di quanto non avessero fatto i suoi colleghi di Amburgo.

Wegener morì nel 1930 mentre guidava una terza spedizione in Groenlandia, probabilmente di un attacco cardiaco causato da sovraffaticamento. I suoi necrologi elogiativi dicono che egli aveva ottenuto una notevole considerazione sia come meteorologo che come esploratore artico; altre fonti aggiungono che egli fu inoltre un organizzatore e amministratore capace e un insegnante lucido e stimolante. La sua opera sulla deriva dei continenti, che sarà certamente la sua eredità permanente, era rimasta di interesse marginale, sebbene l'avesse assorbito profondamente.

Non è certo come Wegener concepisse per la prima volta l'idea che i continenti possono muoversi. Secondo un racconto di dubbia autenticità ne ebbe l'ispirazione durante un viaggio in Groenlandia, mentre osservava la fessurazione di un ghiacciaio (il processo che dà origine agli iceberg). Dai suoi scritti e da quelli dei suoi contemporanei, tuttavia, sembra più probabile che egli sia giunto alla teoria nella stessa maniera in cui vi erano giunti i suoi predecessori: notando su una carta geografica la complementarietà delle linee di costa dell'oceano Atlantico. Secondo lo stesso Wegener l'idea gli venne per la prima volta nel 1910, ma un contemporaneo che lo aveva conosciuto da studente sostiene che egli si era interessato a questo argomento fin dal 1903. Sia che questa idea sia stata maturata per un decennio o per soli due anni, essa fu presentata pubblicamente per la prima volta nel gennaio 1912 in una conferenza alla Associazione geologica tedesca a Francoforte sul Meno. Le prime relazioni pubblicate comparvero più tardi, nello stesso anno, su due riviste tedesche.

Nel 1912 le teorie prevalenti sulla formazione e l'evoluzione della Terra non potevano adattarsi alla deriva dei continenti. I geologi e i geofisici credevano allora che la Terra si fosse formata allo stato fuso e che si stesse ancora solidificando e contraendo. Durante questo processo gli elementi pesanti, quali il ferro, sarebbero sprofondati nel nucleo e i più leggeri, quali il silicio e l'alluminio, sarebbero risaliti alla superficie per dar luogo a una crosta rigida.

Alla maggior parte dei geologi dell'epoca sembrava che questo modello riuscisse molto bene a spiegare i tratti principali della superficie terrestre. Le catene montuose erano prodotte per compressione alla superficie durante la contrazione, proprio come le rughe che si formano sulla buccia di una mela che si asciuga e si contrae. Su scala più vasta la pressione causata dalla compressione, applicata a grandi strutture arcuate, causava il crollo e la subsidenza di alcune regioni alla superficie, creando così i bacini oceanici, mentre altre zone seguitavano a emergere come continenti. I movimenti verticali della crosta venivano considerati interamente plausibili, mentre venivano esclusi i movimenti paralleli alla superficie. Così i continenti e i bacini oceanici erano, alla lunga, intercambiabili; alcune aree continentali sprofondavano più rapidamente delle terre adiacenti e venivano inondate dal mare; contemporaneamente alcune parti del fondo oceanico si sollevavano fino a diventare terre emerse.

La somiglianza o l'identità di numerose piante e animali fossili su continenti lontani veniva spiegata ammettendo l'esistenza di ponti di terraferma che avevano collegato le masse continentali ed erano poi affondati fino al livello del fondo oceanico. La stratificazione dei depositi sedimentari parlava di successive trasgressioni e regressioni del mare sui continenti. Le regressioni potevano venire attribuite alla subsidenza dei bacini oceanici e le trasgressioni al parziale riempimento dei bacini con sedimenti erosi dai continenti. Più o meno al tempo in cui Wegener elaborava la sua teoria della deriva dei continenti venne proposta un'elaborazione del punto di vista tradizionale, secondo la quale i movimenti verticali della crosta sono regolati dall'isostasia: il concetto è che tutti gli elementi del sistema sono in equilibrio idrodinamico. Così i continenti, essendo meno densi dello strato sottostante, galleggiano più in alto dei fondi oceanici.

Wegener individuò un certo numero di difetti e contraddizioni nel modello della Terra che si contrae. Inoltre parecchie caratteristiche importanti della superficie terrestre non venivano affatto spiegate dal modello, a meno che non venissero considerate come prodotte da coincidenze. La più ovvia di queste caratteristiche è la corrispondenza delle coste atlantiche dell'Africa e del Sudamerica. (Nel disegnare questa corrispondenza Wegener non adoperò la vera e propria linea di costa, ma il bordo della piattaforma continentale, che è un limite più significativo; nelle ricostruzioni moderne si usa lo stesso metodo.) Un'altra anomalia si ha nella distribuzione delle catene montuose, che sostanzialmente si limitano a cinture strette e curvilinee, mentre se si fossero prodotte per la contrazione del globo dovrebbero essere distribuite uniformemente su tutta la sua superficie, proprio come le rughe su una mela secca.

Egli scoprì un'altra caratteristica singolare dall'analisi statistica della topografia terrestre. In base ai calcoli dell'area totale della superficie terrestre, Wegener trovò che una grande percentuale della crosta terrestre è a due diversi livelli. Uno corrisponde alla superficie dei continenti e l'altro al fondo dei mari (*si vedano le illustrazioni alle pagine 80 e 81*). È logico aspettarsi una distribuzione simile se la crosta è composta di due strati, quello superiore di rocce più leggere, quali i graniti, e quello inferiore di basalto, gabbro o peridotite, che dovrebbero formare anche i fondi oceanici. Questa interpretazione è rafforzata anche dalla misura delle variazioni locali della gravità terrestre. Non è invece compatibile con un modello della crosta in cui le variazioni in elevazione derivano da sollevamenti e subsidenze casuali; in questo caso ci si dovrebbe aspettare una gaussiana, cioè una distribuzione delle elevazioni a forma di campana centrata su un singolo valor medio del livello.

Wegener trovò sostegno per le sue idee anche nei fossili e in certe caratteristiche geologiche tipiche che sembravano tagliare i margini continentali. Fra tutta la documentazione fossile i rettili forniscono un ottimo esempio. In Brasile e in Sudafrica, e in nessun altro luogo al mondo, si trovano dei fossili di *Mesosaurus*, un piccolo rettile che visse verso la fine dell'era paleozoica, circa 270 milioni di anni fa (si veda *I fossili e la deriva dei continenti* di A. Hallam, in « Le Scienze », n. 54, febbraio 1973). Questa caratteristica distribuzione veniva spiegata tradizionalmente assumendo che un ponte di terraferma aveva collegato i due continenti ed era poi sprofondato. Wegener rifiutò questa spiegazione su basi geofisiche; essa violava infatti il principio dell'isostasia, poiché il materiale del ponte doveva essere meno denso di quello del fondo oceanico e quindi non poteva affondarvi. L'unica alternativa ragionevole era che i continenti un tempo fossero stati uniti e si fossero quindi allontanati.

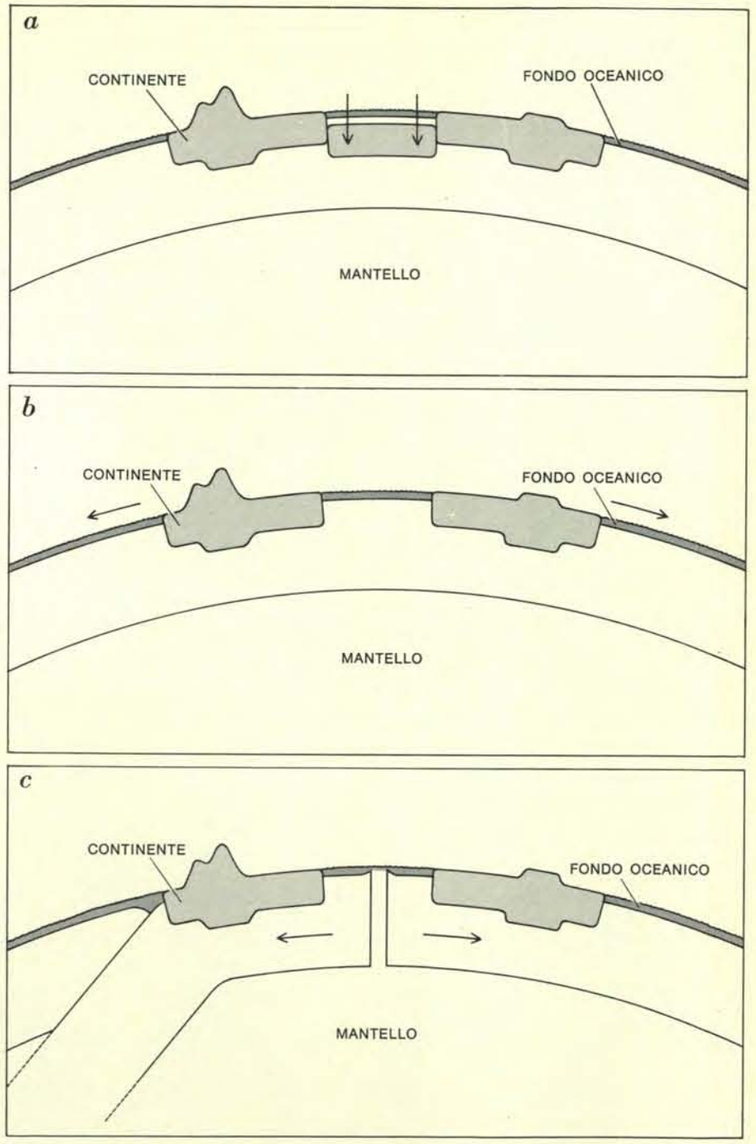
L'evidenza geologica è di natura simile. Per esempio sia in Africa che in Sudamerica, si trovano grandi blocchi di rocce particolarmente antiche; se si riportano insieme i continenti nella posizione giusta i blocchi coincidono perfettamente (*si veda l'illustrazione nella pagina 79*). Lo stesso Wegener riconobbe la portata della scoperta e così la descrisse: « È come se dovessimo rimettere insieme i pezzi strappati di un giornale in base ai loro contorni e poi controllare se le righe di stampa dei pezzi vicini si collegano. Se è così non resta che concludere che i pezzi erano effettivamente uniti in questo modo. Se avessimo a disposizione soltanto una riga per eseguire questo controllo, avremmo ancora una buona probabilità che la ricostruzione sia valida, ma se si hanno *n* righe, questa probabilità viene elevata alla potenza *n*-esima ».

Dobbiamo ancora accennare ad altri dati sperimentali su cui Wegener si basò. Alcune osservazioni geodetiche eseguite agli inizi del secolo XX sembrarono indicare che la Groenlandia si muoveva verso ovest, allontanandosi dall'Europa a una velocità misurabile. Tale movimento avrebbe potuto costituire una convalida diretta della deriva dei continenti, ma non è stato confermato da misure recenti che impiegavano tecniche più accurate.

Allo scopo di risolvere queste contraddizioni Wegener formulò una teoria generale sull'origine dei continenti. Nella sua ricostruzione della storia terrestre tutte le terre emerse del globo erano unite in origine in un unico supercontinente primordiale, che egli chiamò Pangea (*si veda l'illustrazione a pagina 75*). Pubblicò i suoi dati e le sue conclusioni nel 1915, in un libro intitolato *Die Entstehung der Kontinente und Ozeane* (*L'origine dei continenti e degli oceani*).

Le basi geofisiche della teoria di Wegener erano in stretto rapporto col principio dell'isostasia. Ambedue assumono infatti che il substrato sottostante ai continenti si comporta come un fluido altamente viscoso; Wegener inoltre suppose che se una massa rocciosa si può muovere verticalmente in questo fluido, essa dovrebbe essere in grado di muoversi anche orizzontalmente, ammettendo soltanto che venga applicata una forza abbastanza grande. A dimostrazione che tali forze esistono egli citò compressioni orizzontali degli strati piegati nelle catene montuose. C'è un'altra prova elegante della natura fluida del materiale sottostante: la Terra è una sfera schiacciata, che si gonfia leggermente all'equatore, e l'entità del rigonfiamento è esattamente quella che ci si aspetterebbe per una sfera di un fluido perfetto ruotante alla stessa velocità. Essa è quindi un fluido di natura particolare: a sforzi rapidi, come quelli di un terremoto, reagisce come un solido elastico, mentre le sue caratteristiche fluide si possono osservare sui periodi molto più lunghi del tempo geologico. Il suo comportamento è analogo a quello della pece, un materiale che si spezza sotto un colpo di martello, ma fluisce plasticamente per il suo stesso peso, e cioè sotto la forza di gravità, più debole, ma persistente.

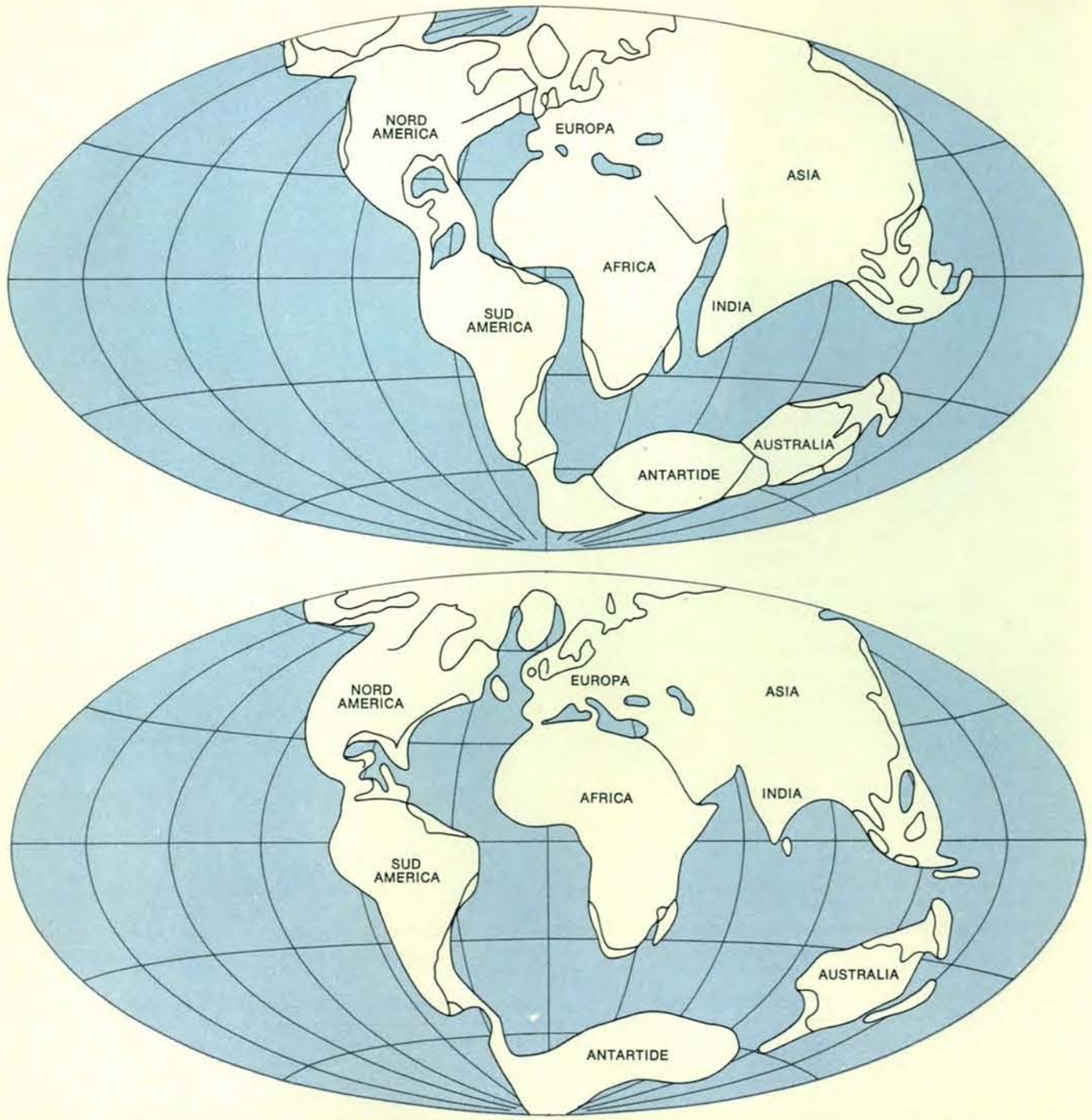
Dopo la prima guerra mondiale Wegener riunì i suoi due principali interessi studiando, insieme a Köppen, i



Wegener sfidò la teoria della struttura terrestre sostenuta dalla maggioranza dei geologi all'inizio del XX secolo. Secondo le vedute tradizionali (*a*) i continenti erano fissi al loro posto e i movimenti laterali erano impossibili. In certe versioni della teoria i moti verticali della crosta, come lo sprofondamento dei ponti di terraferma, erano consentiti, in altre venivano rifiutati in quanto cozzavano contro il principio dell'isostasia, secondo il quale i continenti galleggiano in equilibrio idrostatico su un substrato di materiale più denso. Secondo l'ipotesi di Wegener (*b*) i continenti migrano galleggiando sul substrato, che si comporta come un fluido molto viscoso. Egli suggerì che essi vengono mossi da forze connesse alla rotazione terrestre, ma quest'idea venne presto aspramente criticata e abbandonata. Secondo la teoria moderna (*c*), i continenti vengono trasportati passivamente come passeggeri su grandi zolle rigide. Le zolle vengono allontanate l'una dall'altra là dove sgorga nuovo materiale dal mantello che forma poi nuovo fondo oceanico.

cambiamenti nel clima mondiale attraverso i tempi geologici. Cartografando la distribuzione di certe specie di rocce sedimentarie, egli fu in grado di ricavare la posizione dei poli e dell'equatore nei tempi passati. I risultati più impressionanti sono quelli che egli ottenne per i periodi Permico e Carbonifero, circa 300 milioni di anni fa (*si veda l'illustrazione a pagina 82*). La posizione del polo Sud veniva determinata dalla disposizione dei letti di argille non stratificate, detti *tilliti*, che si sono formati durante il movimento dei ghiacciai. Nella ricostruzione di Wegener della Pangea il polo sud era appena a est dell'attuale Sudafrica e all'interno dell'antica Antartide.

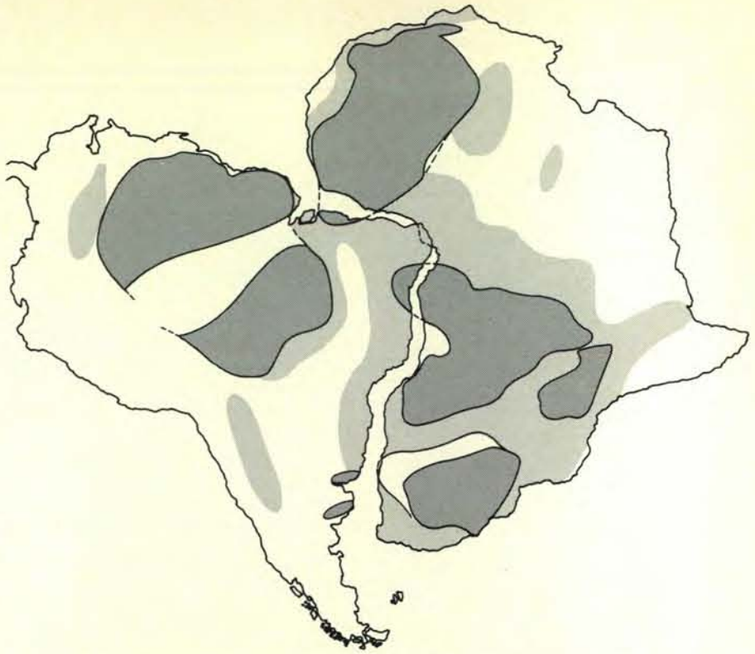
A novanta gradi dal polo Wegener trovò abbondanti prove di una zona equatoriale umida. Queste prove consistono in vasti depositi di carbone che si estendono dalla parte orientale degli Stati Uniti alla Cina; le piante fossili che si possono identificare nei carboni sono di tipo tropicale. Altri indicatori climatici fra le rocce sedimentarie sono il sale e le sabbie deposte dai venti, che suggeriscono la presenza di antichi deserti. In passato come oggi i deserti si formavano specialmente nelle zone dei venti costanti ai due lati dell'equatore. Nell'attuale emisfero settentrionale i carboni del periodo Carbonifero sono seguiti, in strati più recenti, da depositi di sale e dune sabbiose, che Wegener interpretò come segno di una significativa migrazione a sud della posizione dell'equatore. Questo movimento fu confermato da un corrispondente spostamento a sud-est del centro principale dei depositi di tilliti, il che significa che anche il polo si era spostato in quella direzione.



La rottura del Pangea e la formazione dei continenti moderni venne descritta da Wegener in due fasi: i continenti restavano tutti vicini fino all'Eocene, circa cinquanta milioni di anni fa (*carta superiore*). Ancora all'inizio dell'era quaternaria, circa due milioni di anni fa, il Sudamerica e l'Antartide erano collegati da un istmo (*carta inferiore*. I ricercatori moderni hanno riveduto notevolmente la sequenza temporale di Wegener. La rottura in realtà ebbe inizio circa duecento milioni di anni fa.

L'ultima edizione de *L'origine dei continenti e degli oceani* dedicava un capitolo agli antichi climi e conteneva una estesa discussione della deriva polare. Il moto dei poli è un fenomeno completamente diverso dalla deriva dei continenti, ma è impossibile interpretare coerentemente la distribuzione delle tilliti e dei depositi di carbone e di sabbia a meno che i continenti vengano rimessi insieme più o meno come aveva proposto Wegener.

L'ultima edizione contiene anche una documentazione più completa, rispetto alle edizioni precedenti, sulle somiglianze nella geologia dei continenti meridionali, e ciò era dovuto a un produttivo scambio di idee fra Wegener e il geologo sudafricano Alexander L. du Toit. Gli argomenti geofisici fondamentali, però, rimangono notevolmente simili a quelli proposti da Wegener nel suo primo lavoro, scritto quasi vent'anni prima.



Particolari formazioni rocciose, che formano fasce continue dal Sudamerica all'Africa accostate, forniscono l'evidenza geologica che i continenti formavano un tempo un'unica massa. Le zone in grigio scuro sono antichi blocchi (« cratoni »); le zone in grigio chiaro sono aree con rocce un po' più giovani. Wegener considerò che delle continuità come queste sono delle buone prove della teoria della deriva dei continenti.

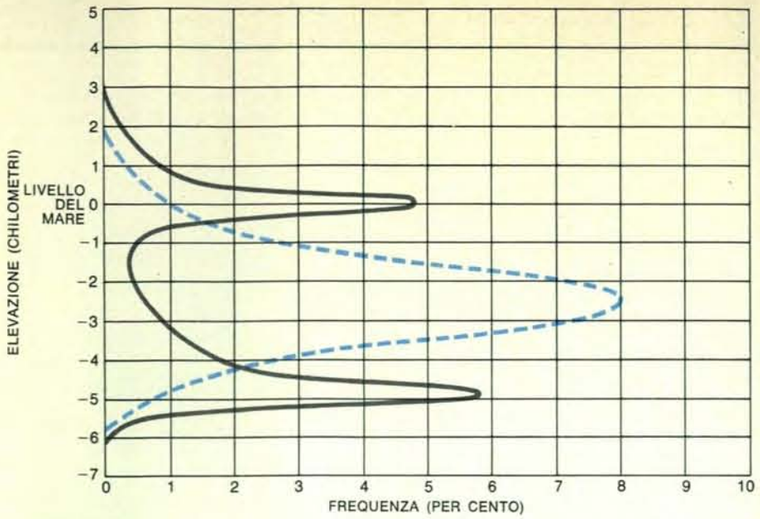
La reazione iniziale del mondo scientifico all'ipotesi di Wegener non fu uniformemente ostile, ma nel migliore dei casi fu mista. Alla sua prima conferenza a Francoforte sul Meno alcuni geologi furono indignati; a Marburg, pochi giorni dopo, sembra che invece l'auditorio fosse più favorevole. In seguito alle prime pubblicazioni parecchi famosi geologi tedeschi dichiararono la loro opposizione allo « spostamento dei continenti » (traduzione più accurata di « deriva » del termine usato da Wegener, *Verschiebung*). Un certo numero di geofisici, d'altra parte, espressero la loro approvazione al concetto. In effetti nel 1921 Wegener poteva affermare di non conoscere alcun geofisico che si opponesse all'ipotesi della deriva.

Non sembra che le prime pubblicazioni, compresa la prima edizione de *L'origine dei continenti e degli oceani*, siano state lette molto fuori della Germania: fu soltanto quando la terza edizione, pubblicata nel 1922, fu tradotta in diverse altre lingue, compreso l'inglese, due anni dopo, che l'ipotesi di Wegener trovò un uditorio internazionale. Come in Germania, la sua opera ebbe inizialmente un'accoglienza abbastanza buona o almeno non venne lasciata cadere completamente. Secondo i rendiconti pubblicati, a un convegno della Associazione britannica per l'avanzamento della scienza tenuto nel 1922, la discussione sulla deriva dei continenti fu « viva, ma inconcludente ». Come è facile aspettarsi, molti erano scettici, ma l'accoglienza generale riservata alla teoria era favorevole. Più o meno in quel momento parecchi celebri geologi di ambedue le sponde dell'Atlantico si dichiararono favorevoli alla teoria.

L'antagonismo irragionevole e ostinato alle idee di Wegener, che doveva diventare l'ortodossia della geofisica fino agli anni sessanta, iniziò a svilupparsi verso la metà degli anni venti. Due eventi provocarono questo inasprimento delle resistenze: uno fu la pubblicazione di un trattato intitolato *La Terra*, di Harold Jeffreys dell'Università di Cambridge; e l'altro fu un simposio dell'Associazione americana dei geologi del petrolio, tenuto nel 1928.

Jeffreys attaccò la teoria di Wegener in quello che era forse il suo punto più debole: la natura delle forze a cui Wegener attribuiva il moto dei continenti. Wegener aveva scritto che la deriva della Americhe verso ovest si poteva spiegare come conseguenza delle maree nella crosta terrestre: egli aveva chiamato *Polflucht*, cioè « fuga dai poli », la forza responsabile della migrazione dell'India verso il nord e nelle catene montuose alpina e himalayana. Jeffreys fu in grado di dimostrare mediante semplici calcoli che la Terra è di gran lunga troppo resistente perché tali forze possano deformarla anche leggermente. Se così non fosse le catene montuose crollerebbero sotto il loro stesso peso e il fondo oceanico sarebbe perfettamente piatto. Se le forze di marea fossero abbastanza intense da muovere i contenenti verso ovest, esse sarebbero anche abbastanza grandi da fermare la rotazione terrestre nel giro di un anno.

Queste obiezioni sono decisive, e il meccanismo proposto da Wegener è stato abbandonato da lungo tempo. (La teoria moderna attribuisce il movimento dell'espansione dai fondi oceanici lungo un sistema di dorsali medio oceaniche, dove rocce fuse sgorgano dall'interno della Terra.) Oggi tuttavia ci rendiamo conto che Jeffreys non aveva fatto altro che dimostrare l'inadeguatezza delle ipotetiche forze motrici, senza rifiutare la teoria; aveva semplicemente ignorato la maggior parte dei dati empirici che costituivano la parte più sostanziale degli argomenti di Wegener. Jeffreys aveva liquidato la deriva dei poli sostenendo che era geofisicamente impossibile.



L'analisi statistica delle elevazioni della superficie terrestre fornì a Wegener un argomento a sostegno della sua ipotesi. Se le caratteristiche topografiche si sono formate per sollevamenti e subsidenze casuali, ci si aspetterebbe una distribuzione casuale (gaussiana) centrata sulla elevazione media (*curva in colore*). Invece si hanno principalmente due livelli preferenziali (*curva in grigio*). Uno, prossimo al livello dei mari, rappresenta l'elevazione media delle piattaforme continentali, l'altro quella dei fondi oceanici.

Fra i partecipanti al simposio dell'Associazione americana dei geologi del petrolio, la maggioranza era ostile, in vari gradi, alla teoria di Wegener; solo uno era decisamente favorevole. I rendiconti del simposio sono principalmente un coro di critiche. I relatori sostenevano che l'apparente coincidenza dei profili dei continenti atlantici era imprecisa e non teneva conto dei movimenti verticali della crosta. Le formazioni rocciose simili sulle due sponde dell'oceano, dopo tutto, non erano in stretti rapporti, e in ogni caso le somiglianze attuali non implicavano necessariamente la contiguità nel passato. Gli antichi animali potevano essersi spostati attraverso ponti di terraferma. Le tilliti del Carbonifero e del Permico del Sudafrica e di altre zone probabilmente non erano glaciali e probabilmente i carboni dell'emisfero settentrionale non erano tropicali; i dati per il moto della Groenlandia non erano decisivi.

Altri oratori pretesero che la teoria risolvesse dei paradossi che essa ha affrontato soltanto nelle sue versioni più moderne. Essi chiedevano per esempio perché, se il continente americano si può muovere lateralmente spostando i fondi oceanici, esso si è arricciato sul bordo occidentale, formando le catene montuose delle cordigliere. La forza compressiva che aveva formato queste montagne non suggeriva forse una notevole resistenza da parte del fondo oceanico, supposto fluido? E ancora, perché il Pangea è rimasto intatto per la maggior parte della storia della Terra, e poi si è frantumato improvvisamente nel giro di poche decine di milioni di anni?

Infine, oltre a mettere in dubbio l'interpretazione e le conclusioni di Wegener, alcuni partecipanti al simposio attaccarono violentemente le sue credenziali e i suoi metodi. Si comportava come un avvocato, sostenevano, che sceglie fra i fatti da presentare solo quelli che appoggiano la sua ipotesi. Egli « si prendeva delle libertà col nostro globo », « faceva un gioco senza regole restrittive, in cui non era tracciato alcun codice di condotta ».

Wegener non partì in battaglia per difendere le sue teorie da queste critiche. Uomo ormai di mezza età, egli disponeva di un tempo limitato per la ricerca e non si sentì in grado di tener testa alla crescente ondata di pubblicazioni; si contentò di lasciare il campo a ricercatori più giovani.

Dopo la morte di Wegener i geologi e i geofisici divennero ancora più ostili alla sua ricostruzione della storia della Terra. Negli Stati Uniti la reazione fu particolarmente violenta; per un geologo americano esprimere simpatia per l'idea della deriva dei continenti significava giocarsi la carriera.

Per ironia della sorte, non appena fu emesso il verdetto di condanna, la teoria fu notevolmente rafforzata dai contributi di du Toit e di Arthur Holmes, dell'Università di Edimburgo. Essi eliminarono alcuni degli argomenti più deboli di Wegener, misero insieme altri dati a sostegno della teoria ed elaborarono un meccanismo più plausibile per il movimento. Holmes suggerì che i continenti sono mossi dalle correnti convettive nel mantello terrestre. La sua ipotesi non era completamente soddisfacente, ma superava con successo le critiche di Jeffreys e dei suoi seguaci. Inoltre anticipava la spiegazione moderna del perché i continenti si muovono.

A posteriori, sapendo quanti aspetti della teoria di Wegener sono stati confermati negli ultimi vent'anni, siamo in grado di apprezzare immediatamente i suoi risultati. Egli esaminava criticamente un modello della Terra accettato quasi universalmente. Quando vi scoprì delle debolezze e delle incoerenze ebbe tanto coraggio e spirito d'indipendenza da abbracciare un'alternativa radicale. Inoltre le sue conoscenze erano così vaste da permettergli di andare in cerca e di valutare criticamente nuovi dati di numerose discipline. Egli applicò le stesse qualità mentali alla spiegazione degli antichi climi e della deriva dei poli.

Il rigore intellettuale che Wegener portò nella sua opera è illustrato dai suoi stessi scritti e testimoniato dalle parole di coloro che lo conobbero bene. In una lettera a Köppen scritta nel 1911 Wegener difendeva i suoi punti di vista sulla deriva dei continenti; la lettera è riprodotta nella biografia di Wegener scritta dalla sua vedova.

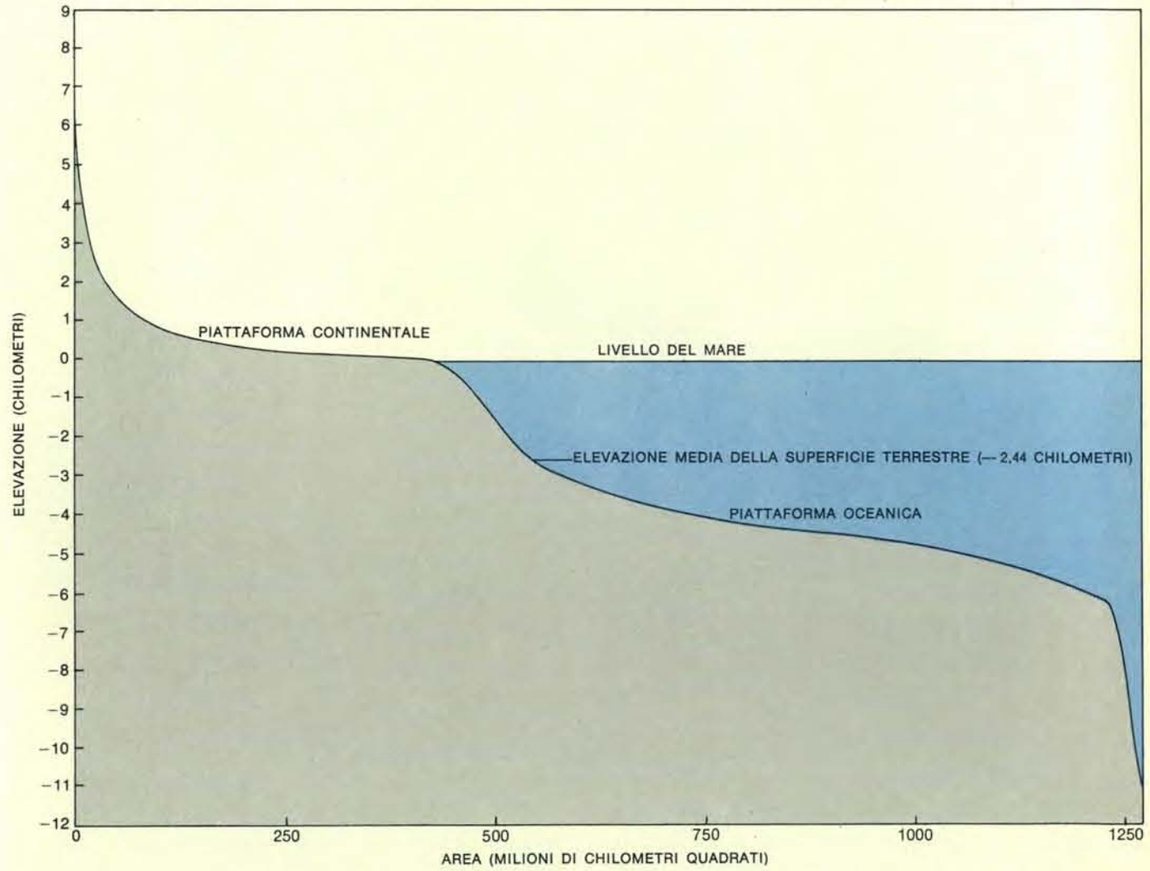
« Lei considera il continente primordiale come un prodotto della mia fantasia, ma è solo un problema di interpretazione dei dati. Io sono giunto a questa idea in base al fatto che le linee di costa combaciano, ma le prove devono venire da osservazioni geologiche. Esse ci obbligano ad ammettere, per esempio, un legame di terre emerse fra il Sudamerica e l'Africa. Questo si può spiegare in due modi: lo sprofondamento di un continente, che li collegava, o la separazione. In precedenza, per il principio di permanenza, non dimostrato, è stato considerato soltanto il primo, e la seconda possibilità è stata ignorata; ma la moderna teoria dell'isostasia e, più in generale, le nostre idee geofisiche correnti, si oppongono allo sprofondamento di un continente, perché esso è più leggero del materiale su cui poggia. Perciò siamo obbligati a considerare l'interpretazione alternativa. E se ora troviamo molte semplificazioni sorprendenti, e se cominciamo almeno a dare un senso reale a un'intera messe di dati geologici, cosa dovremmo aspettare per gettar via i vecchi concetti? Forse ciò è rivoluzionario? Non credo che le vecchie idee abbiano più di un decennio di vita. Attualmente la nozione di isostasia non è ancora stata elaborata completamente; quando lo sarà, le contraddizioni implicite nel vecchio modello appariranno in piena luce ». Ciò, evidentemente, sembrava molto ovvio a Wegener, ma è chiaro che egli sottovalutava la tenacia di chi restava ancorato ai vecchi concetti.

Il successo di Wegener nel costruire una teoria sistematica della storia della Terra partendo da osservazioni sparse e apparentemente non collegate potrebbe essere attribuito alla larghezza nel modo di affrontare il problema, e forse anche al fatto che egli non era uno specialista. Si può avere un chiarimento sul suo metodo nell'affrontare i problemi scientifici leggendo un necrologio scritto da Hans Benndorf, professore di fisica e collega di Wegener a Graz. « Wegener acquisiva le sue conoscenze principalmente in un modo intuitivo, mai, o assai di rado, deducendole da formule e in questo caso esse dovevano essere molto semplici. Se poi aveva a che fare con problemi riguardanti la fisica, cioè un campo lontano da quello di cui era esperto, rimanevo spesso sbalordito dalla precisione dei suoi giudizi. Con che facilità, con quale senso dei punti importanti egli si districava attraverso i più complicati lavori dei teorici! Spesso, dopo una lunga pausa di riflessione, diceva: "Io credo questo e questo", e la maggior parte delle volte aveva ragione, come noi avremmo stabilito parecchi giorni dopo, in seguito a un'analisi rigorosa. Wegener possedeva un senso per le cose significative che sbagliava di rado».

Le affermazioni di Benndorf sul metodo di Wegener sono confermate dalle affermazioni di Wilhelm Max Wundt, il quale conobbe Wegener quando era studente a Berlino: « Alfred Wegener si preparava ad affrontare i suoi problemi scientifici con delle capacità normalissime in matematica, fisica e nelle altre scienze naturali. Mai, in tutta la sua vita, egli fu restio ad ammettere questo fatto. Tuttavia egli aveva la capacità di applicare queste doti con grande volontà e cosciente determinazione. Aveva una straordinaria capacità di osservazione, una capacità di capire cosa è contemporaneamente semplice e importante e cosa si può sperare che porti dei risultati. Oltre a ciò possedeva una logica rigorosa, che gli permetteva di mettere giustamente insieme tutto ciò che riguardava le sue idee ».

Se Wegener era realmente lo scienziato abile e intelligente descritto dai suoi contemporanei e se le sue conclusioni erano ben radicate in fatti e ragionamenti, nasce un problema immediato: perché l'opposizione alle sue idee fu così forte, vasta e persistente?

Una spiegazione possibile è che la

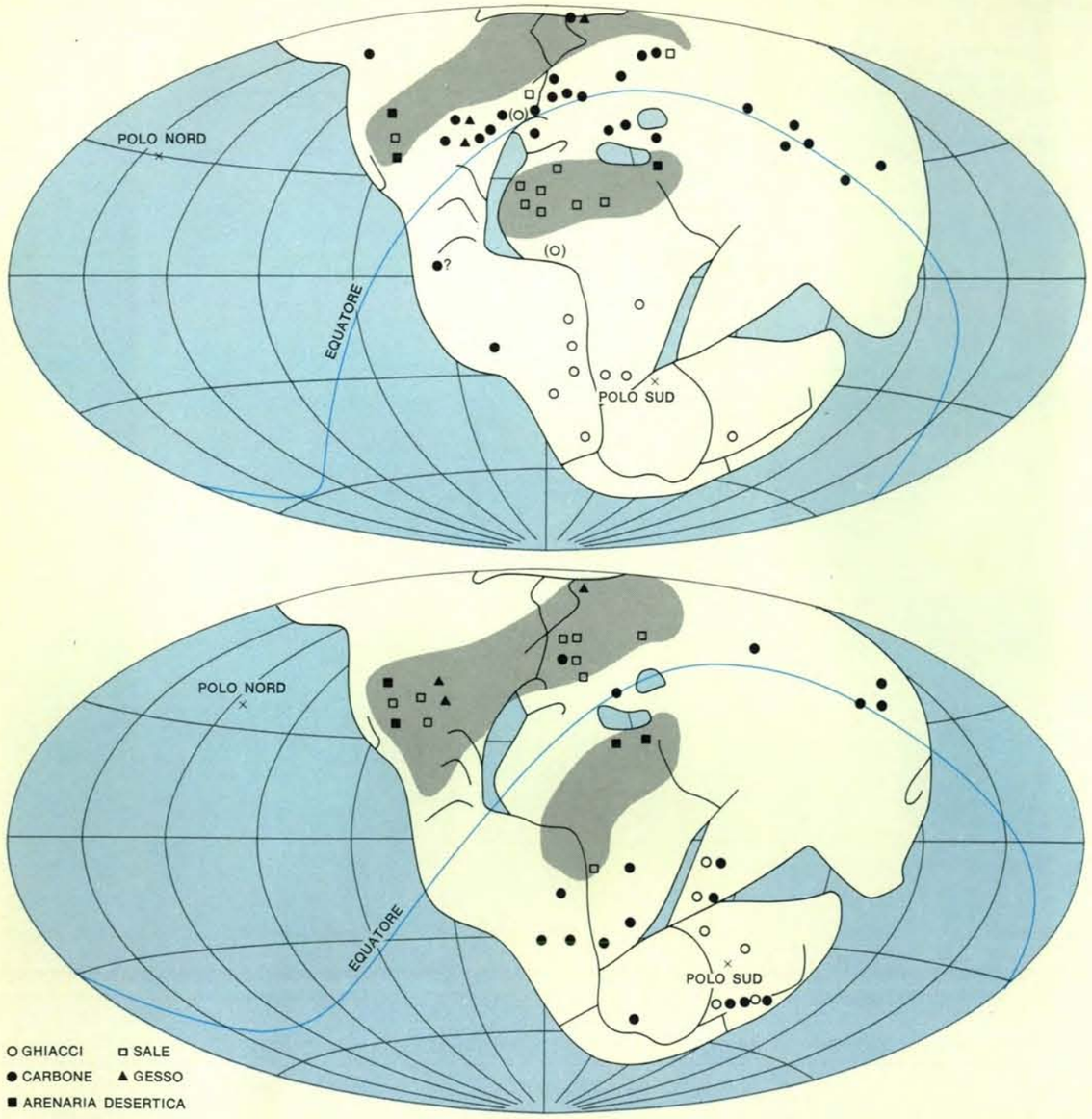


La distribuzione delle elevazioni, se viene correlata all'area della superficie terrestre, conferma che la crosta ha due livelli fondamentali. Circa 800 milioni di chilometri quadrati, sui circa 1300 milioni della superficie terrestre, appartengono a piattaforme continentali o piattaforme oceaniche. Questa distribuzione è in accordo col modello fornito dalla teoria di Wegener.

teoria di Wegener era «prematura» per i tempi in cui egli la presentò. Gunther S. Stent, dell'Università di Berkeley, in California, sostiene che un'idea va considerata prematura se non può essere collegata da una serie di passi semplici e logici alla conoscenza canonica, cioè generalmente accettata, di quel tempo. Un principio simile, espresso da Michael Polanyi, uno studioso inglese di filosofia e sociologia della scienza, sostiene che nella scienza ci dev'essere sempre un'opinione prevalente, sulla natura delle cose, alla quale va confrontata la validità di tutte le asserzioni. Ogni osservazione che sembra contraddire il punto di vista ben stabilito sulle cose va preliminarmente giudicata non valida e accantonata, nella speranza che in seguito essa si riveli falsa o irrilevante. Questa interpretazione di come procede la scienza suggerisce che i geologi e i geofisici dovevano venire sommersi dai dati, come avvenne negli anni sessanta, prima che si decidessero ad abbandonare la dottrina ben stabilita dei continenti stazionari. Le innovazioni di Wegener, proprio perché erano innovatrici, dovevano essere ibernate finché da esse non potesse sorgere una nuova ortodossia.

Quasi certamente c'è del vero in questa interpretazione, e bisogna convincersi che il rifiuto dell'opera di Wegener era una circostanza necessaria al progresso ordinato della scienza; ciononostante questa analisi non è del tutto convincente. Potrebbe spiegare l'indifferenza alla teoria della deriva dei continenti, ma non l'atteggiamento di tanti scienziati che la relegarono nel regno della fantasia. E non spiega neppure perché il modello tradizionale della Terra venne conservato anche dopo che Wegener ebbe dimostrato che vi erano delle contraddizioni, nonostante queste contraddizioni non venissero mai risolte. I paleontologi continuarono a basarsi su misteriosi ponti di terraferma, per esempio, mentre contemporaneamente i geofisici, che avevano adottato il principio dell'isostasia, insistevano che lo sprofondamento di tali ponti era impossibile. È stato suggerito che l'ostacolo principale che impediva che la deriva dei continenti venisse accettata era la mancanza di una forza motrice plausibile, dopo che Jeffreys aveva rifiutato quella proposta inizialmente da Wegener. Eppure, se è così, perché la teoria di Holmes delle correnti convettive ricevette così poca considerazione? Del resto la natura del meccanismo che sposta i continenti rimane incerta ancora oggi, eppure la tettonica a zolle si è così bene affermata che coloro che ne rifiutano i principi fondamentali vengono di solito trattati da reazionari.

Forse la lunga gestazione della teoria di Wegener può venire spiegata meglio come una conseguenza dell'inerzia. Al simposio del 1928 dell'Associazione americana dei geologi del petrolio sembra che un geologo abbia detto: «Se dobbiamo credere all'ipotesi di Wegener, dobbiamo dimenticare tutto quello che abbiamo imparato negli ultimi 70 anni e ricominciare tutto da capo». Bisogna ancora ricordare che per i geologi Wegener era un *outsider*; essi devono averlo considerato un dilettante. Oggi, ovviamente, ci rendiamo conto che la sua posizione era per lui un vantaggio, dato che non aveva niente da guadagnare a conservare il punto di vista tradizionale. Inoltre ci rendiamo conto che in fondo egli non era affatto un dilettante, ma un ricercatore interdisciplinare di grande talento e immaginazione, che si è certamente meritato un posto nel pantheon dei grandi scienziati.



La deriva dei poli venne proposta da Wegener per spiegare la distribuzione degli antichi climi. Certi tipi di carbone sono una prova di clima tropicale; le tilliti, che segnalano le glaciazioni di clima polare e i depositi di sale, di gesso e le arenarie desertiche sono una prova di clima arido. I simboli usati per la carta vengono spiegati nella legenda a sinistra; inoltre le zone aride, caratteristiche delle latitudini con venti costanti, sono rappresentate in grigio. La carta superiore mostra la ricostruzione di Wegener per il Carbonifero, circa 300 milioni di anni fa; la carta inferiore è per il periodo Permico, circa 230 milioni di anni fa. (La distorsione dell'equatore è provocata dal tipo di proiezione geografica impiegata.) I movimenti dei continenti e dei poli non sono correlati, ma una ricostruzione coerente si può fare solo rimettendo insieme i continenti come fece Wegener.

BANCO DI NAPOLI

Istituto di Credito di diritto pubblico
Fondato nel 1539
Fondi patrim. e riserve: L. 113.490.924.640

■ TUTTE LE OPERAZIONI E I SERVIZI DI BANCA

Credito Agrario
Credito Fondiario
Credito Industriale
e all'Artigianato
Monte di Credito su Pegno
Servizi di Ricevitorie
Esattorie e Tesorerie

- Direzione Generale in Napoli
- Rappresentanza della Direzione Generale in Roma
- Delegazione del Servizio Estero in Roma
- Oltre 500 Filiali in Italia

- Filiali all'estero: Buenos Aires • New York
- Rappresentanze all'estero: Bruxelles • Buenos Aires • Francoforte s/M Londra • New York • Parigi • Zurigo
- Rappresentanza per la Bulgaria: VITOCHA, Sofia
- Ufficio Cambio permanente a bordo della m/n «Augustus»

CORRISPONDENTI IN TUTTO IL MONDO

I coproliti dell'uomo

Sono state rinvenute feci fossili umane che risalgono a un periodo che va da poche centinaia di anni a oltre 300 000 anni fa. La loro analisi ci fornisce preziose informazioni su dieta, ambiente e abitudini

di Vaughn M. Bryant jr. e Glenna Williams-Dean

I reperti fossili umani sono pochi e di scarso rilievo, ma costituiscono una fonte di informazione sul comportamento che è negata a coloro che studiano, per esempio, i fossili dei dinosauri. Laddove le condizioni sono favorevoli alla conservazione di materiale organico gli archeologi spesso portano alla luce, negli insediamenti preistorici o nelle loro vicinanze, feci umane fossilizzate. I reperti di questo tipo, chiamati dai paleontologi coproliti, non sono infrequenti fra i resti di animali più antichi dell'uomo; alcune rocce sedimentarie, in realtà, sono costituite in larga misura da pallottoline fecali di animali marini. Questi antichi coproliti, tuttavia, sono diventati pietra, il che rende praticamente impossibile analizzarne l'originario contenuto organico. Invece, grazie alle tecniche di laboratorio messe a punto negli ultimi dieci anni, è divenuto oggi possibile effettuare comunemente un'analisi di questo tipo sui coproliti umani. L'analisi dei coproliti che risalgono a periodi che vanno dal paleolitico medio alla tarda preistoria fornisce agli studiosi informazioni notevolmente precise sulle diete e le attività stagionali dei più antichi cacciatori e coltivatori del nuovo e del vecchio mondo.

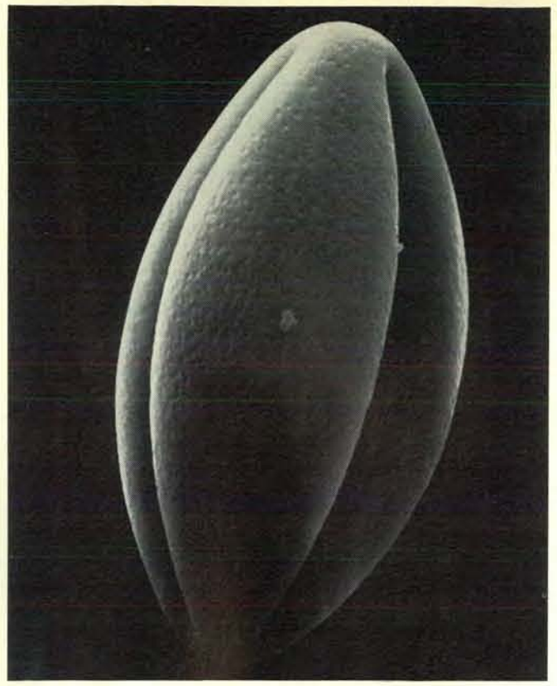
Il ricercatore che effettua lo scavo può non essere in grado di distinguere tra i coproliti dell'uomo e quelli di altri animali, ma l'analisi di laboratorio quale viene oggi praticata permette quasi sempre una corretta identificazione. Il passo determinante del procedimento, introdotto negli anni sessanta dal compianto Eric O. Callen del Macdonald College in Canada, è rappresentato dalla immersione del campione in una soluzione di fosfato trisodico. Nel Laboratorio di ricerche antropologiche dell'Università A. & M. del Texas, dove abbiamo compiuto la maggior parte del nostro lavoro, noi immergiamo il campione per almeno 72 ore. Se il campione proviene da un mammifero onnivoro o carnivoro, il liquido in genere rimane trasparente; se si colora, assume una colorazione marrone chiara. Se il campione deriva da un mammifero erbivoro, il liquido è trasparente e talvolta giallastro. Se il campione è di origine umana, il liquido è opaco e assume un colore marrone scuro o nero. A parte l'uomo, si conosce solo un mammifero, il coati della famiglia dei procionidi, le cui feci danno luogo a una reazione chimica del genere.

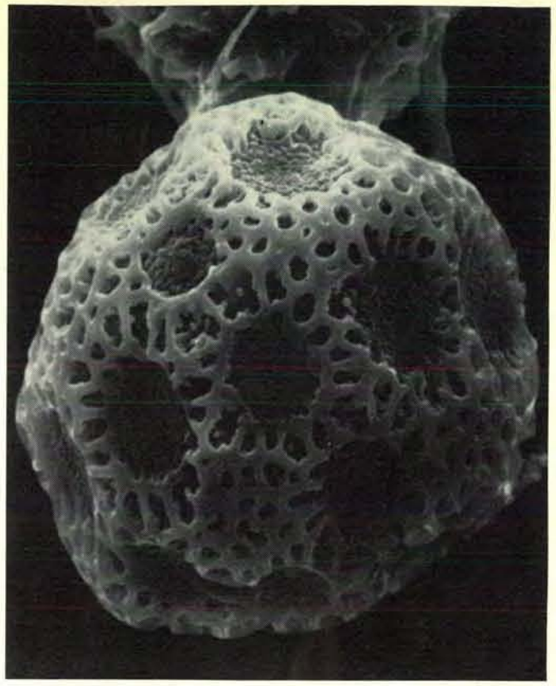
Se i risultati della reazione non sono conclusivi, i passaggi successivi eliminano quasi ogni dubbio. Il campione viene passato attraverso un setaccio che lascia filtrare tutte le particelle di diametro inferiore a 850 micron e poi attraverso un setaccio più fine che lascia passare particelle di diametro inferiore a 150 micron. In tutte e due queste fasi il materiale che residua viene lavato delicatamente con acqua distillata in modo che possano passare attraverso il vaglio anche quelle piccole particelle, come granuli di polline, che aderiscono ai detriti più grossi. I residui che rimangono su entrambi i setacci vengono asciugati e preparati per l'esame microscopico. La frazione liquida viene centrifugata per raccogliere granuli di polline, cristalli vegetali e qualsiasi altra particella essa contenga.

Poiché la dieta umana è tipicamente costituita sia da materiale vegetale sia da materiale animale e poiché gli stesseri umani da almeno mezzo milione di anni hanno imparato a cuocersi il cibo, i residui solidi ottenuti comprendono materiali di origine molto varia. Insieme con pezzi di carbone dovuti al processo di cottura possono esservi fibre vegetali, semi rotti o macinati, peli e frammenti d'osso, gusci di noce o d'uovo, conchiglie di molluschi, penne e rivestimenti chitinosi di insetti. Alcuni di questi materiali possono essere presenti anche in coproliti non umani, ma quando se ne trova una grande varietà il campione è quasi certamente umano. Inoltre, se si considera il contenuto in polline del campione, la possibilità di confusione fra fonte umana e non umana può essere ulteriormente ridotta.

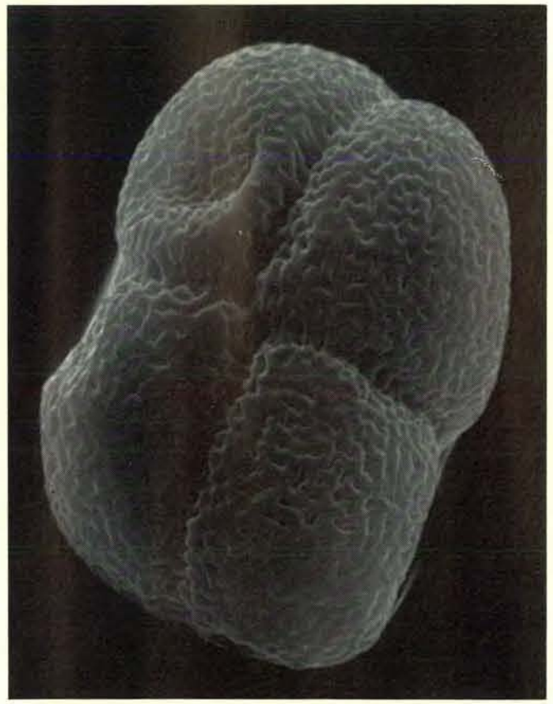
In linea di principio è possibile raccogliere, da tali materiali, informazioni non solo sulle preferenze alimentari di particolari gruppi o individui, ma anche sullo stato di salute dell'individuo (presenza o assenza di parassiti), sulle tecniche di preparazione del cibo, sulle stagioni dell'anno in cui il sito era abitato e sulle condizioni ambientali dell'epoca a cui il campione risale. In pratica la maggior parte delle informazioni concerne le preferenze alimentari e le attività stagionali. Per esempio, l'analisi del polline rivela che circa 2800 anni orsono gli indiani del Texas sudoccidentale raccoglievano e consumavano i fiori di numerose piante di zone desertiche e semidesertiche. Più difficile da stabilire è se tali piante venivano succhiate o masticate per estrarne il nettare, o se ne facevano infusi, o se le mangiavano in insalata; forse venivano consumate in tutti questi diversi modi.

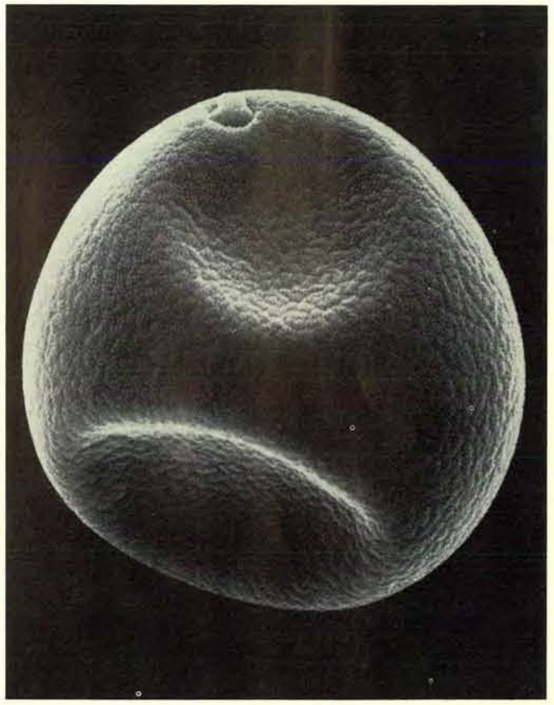
I semi hanno un guscio duro che resiste alla digestione. Un guscio, o anche un semplice frammento di guscio, ha in genere una forma o un tipo di superficie caratteristica che permette di identificare la famiglia di piante, e in alcuni casi anche il genere o la specie da cui proviene. Un esempio di analisi di frammenti di semi che ha gettato luce sulle antiche tecniche di prepara-



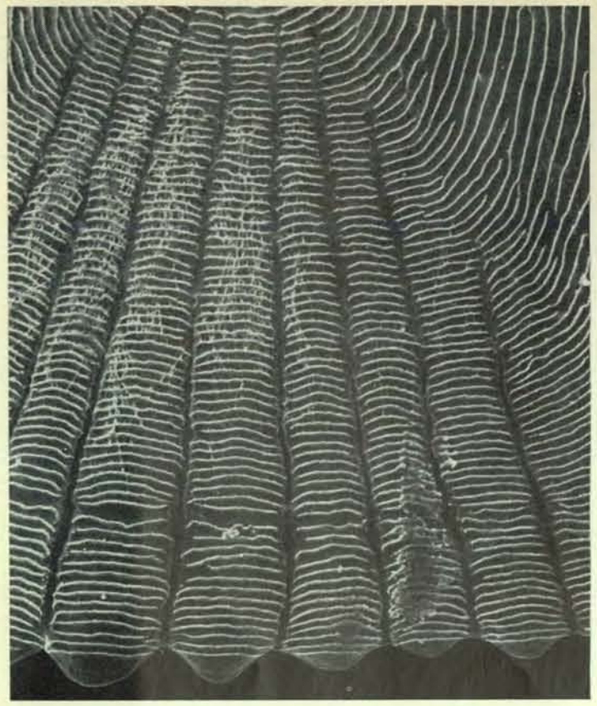


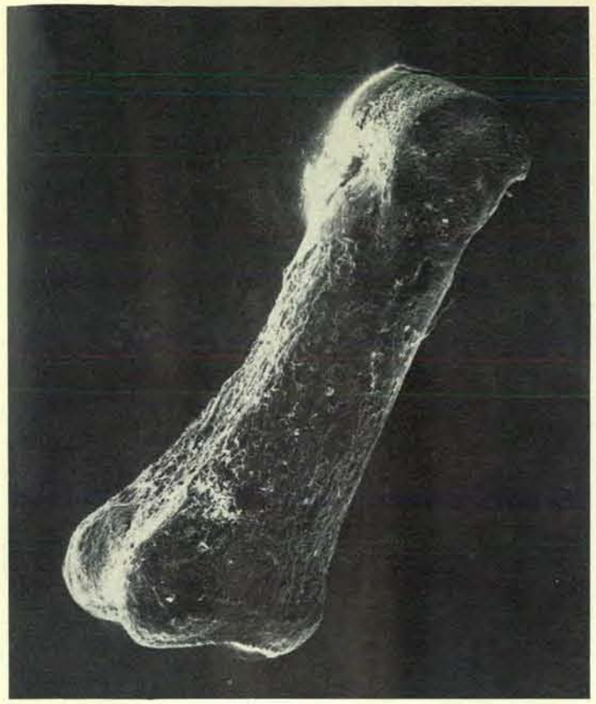
I granuli di polline di mesquite (*Prosopis, a sinistra*) e di fico d'India (*a destra*) sono esempi di polline trasportato da insetti trovato in antichi coproliti. In queste microfotografie elettroniche i pollini risultano ingranditi rispettivamente 3100 e 1350 volte. Entrambi i granuli sono moderni; essi provengono dalla collezione di campioni dell'autore che serve da materiale di confronto.

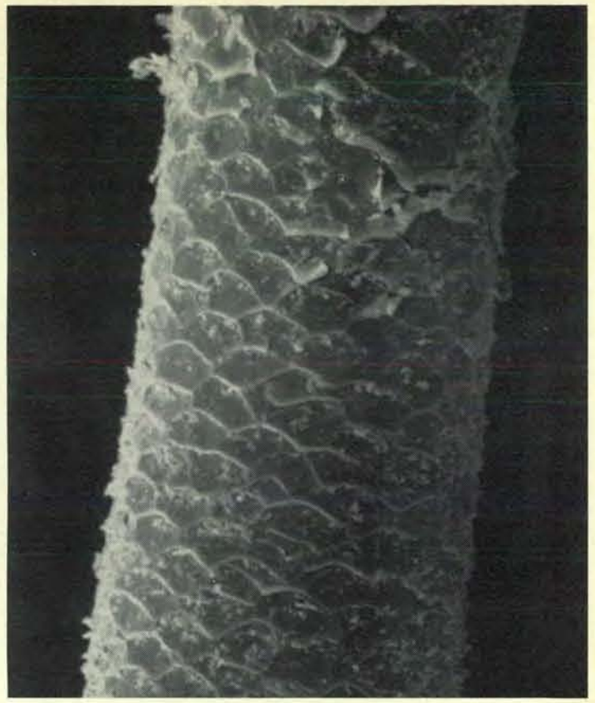




I pollini trasportati dal vento sono qui rappresentati da una tetrade di granuli di polline di *Typha* (*a sinistra*) e da un granulo di sorgo (*a destra*). Essi sono ingranditi rispettivamente 4500 e di 2250 volte. Questo tipo di polline viene diffuso dal vento in determinate stagioni e chiunque prepari del cibo o beva acqua in quei periodi ha molta probabilità di ingerirne.







Le scaglie di pesce, come quelle dei rettili, attraversano il canale alimentare inalterate. Questa scaglia di pesce sole, caratterizzata da un disegno ad anelli concentrici, è ingrandita 175 volte. Anche le penne non sono modificate dal processo digestivo; spesso la trama delle barbe identifica la specie di uccello a cui appartiene. Questo esemplare è una penna di pollo ingrandita 550 volte.

Spesso è possibile identificare piccole ossa, se esse non sono troppo frammentate dalla masticazione. Questo osso di roditore, ingrandito 45 volte, proviene da un riparo di roccia del Texas. I peli degli animali, come le scaglie e le penne resistono alla digestione. Nella microfotografia è visibile un disegno a scaglie che è caratteristico dei ruminanti. L'esemplare è ingrandito 1100 volte.

zione del cibo è rappresentato dallo studio di Callen sui coproliti di Tehuacán, un sito messicano che abbraccia circa 12 000 anni di preistoria del nuovo mondo. Alcuni dei frammenti di guscio di miglio raccolti da Callen rivelavano segni di frammentazione, il che ha dimostrato che erano stati trattati mediante schiacciamento. Altri frammenti, invece, risultarono divisi, segno che i semi non erano stati schiacciati bensì strofinati in su e in giù su un *metate* di pietra.

Anche noi abbiamo ottenuto interessanti informazioni sulle tecniche di preparazione del cibo da uno studio sui coproliti provenienti da un riparo di roccia del Texas sudoccidentale. Alcuni frammenti di guscio di semi di cactus contenuti nei campioni di cinque degli otto strati del sito hanno mostrato che i semi di cactus erano stati trattati, prima di essere mangiati, mediante schiacciamento o macinatura. I gusci di semi e di altre piante, in particolare quelli di miglio e chenopodiacee, erano carbonizzati, segno che non erano stati macinati, ma arrostiti. Molti gusci di semi contenuti nei coproliti non sono rotti; quando un seme non è apprezzato di per sé come cibo, può essere tuttavia ingerito insieme alla frutta. Come risultato è stato possibile dimostrare che la dieta dei cacciatori e coltivatori preistorici di molti siti del nuovo mondo comprendeva elementi come peperoncini rossi, uva, pomodori, guava, more e meloni.

Un tipo più sottile di informazione sulle abitudini alimentari, limitata alle regioni in cui l'acqua del terreno contiene molti minerali in soluzione, è fornito dal fitolito, o cristallo delle piante. A differenza degli animali, le piante non possono eliminare le sostanze inorganiche nei loro fluidi. Il calcio e il silicio che vengono assunti con l'acqua del terreno vengono depositati nei tessuti della pianta in forma cristallina, per la maggior parte sotto forma di sali di calcio e biossido di silicio. La forma di un particolare cristallo non è in genere associata con una specie di pianta. Tuttavia, nell'esaminare i fitoliti di piante della stessa famiglia o genere che crescono nel Texas sudoccidentale, noi abbiamo trovato certi cristalli che hanno forme simili. Per esempio, i cristalli di ossalato di calcio che si osservano in quattro specie di fico d'India (*Opuntia*) sono all'apparenza molto simili. Vi è però un fitolito specifico della specie. Si tratta di un cristallo a forma romboidale che si trova solo nel tessuto di una specie di agave, l'*Agave lecheguilla* (*si veda l'illustrazione a pagina 92*). L'identificazione di questi cristalli ci ha consentito di dimostrare che un certo numero d'individui tra gli indiani arcaici che successivamente visitarono il riparo di roccia nel Texas sudoccidentale avevano sicuramente mangiato fichi d'India e agavi, anche se i coproliti in cui i cristalli erano stati trovati non contenevano fibre, semi o polline identificabili di tali piante.

Naturalmente si ottengono informazioni sia da materiali animali sia da materiali vegetali. Per esempio, in un sito del Nevada, la grotta Lovelock, Lewis K. Napton e O.A. Brunetti dell'Università di California a Berkeley poterono identificare le varie specie di uccelli acquatici mangiati dagli occupanti per mezzo di un'indagine sui frammenti di penne trovati nei coproliti. Gli abitanti si erano nutriti di aironi, svassi, gallinelle d'acqua e oche. Nelle analisi da noi effettuate sul materiale proveniente dal riparo di roccia del Texas, abbiamo identificato vari resti animali, tra cui frammenti di cavallette e altri insetti, spine e squame di pesce e ossa di piccoli mammiferi e rettili (*si veda l'illustrazione a pagina 93*).

Quando la preda era voluminosa, per esempio un cervo o un bisonte, la macellazione e la cottura riducevano al minimo la possibilità di ingerire ossa o frammenti di osso. Quasi inevitabilmente, tuttavia, alcuni peli rimanevano attaccati alla carne e venivano inghiottiti. Abbiamo riconosciuto i peli di un piccolo mammifero, il topo dei campi, in alcuni coproliti del riparo di roccia e abbiamo persino recuperato materiale animale apparentemente deteriorabile come le larve di insetti.

C'è ancora un tipo di materiale animale che percorre tutto il canale alimentare rimanendo in larga misura inalterato: le scaglie dei pesci e dei rettili. Il separare le une dalle altre non presenta grandi difficoltà. Le scaglie di pesce mostrano un disegno ad anelli di crescita concentrici e sono generalmente arrotondate. Le scaglie dei rettili non hanno anelli di crescita e sono spesso appuntite a un'estremità. In circostanze ideali, e disponendo di un'adeguata raccolta di materiale di riferimento che serve da confronto, è possibile sapere qualcosa di più sul reperto e non solo se si tratta di un pesce o di un rettile. Napton e un altro collega di Berkeley, Robert F. Heizer, sono stati capaci di identificare tre diverse specie di pesce di cui si nutrivano gli occupanti della grotta di Lovelock.

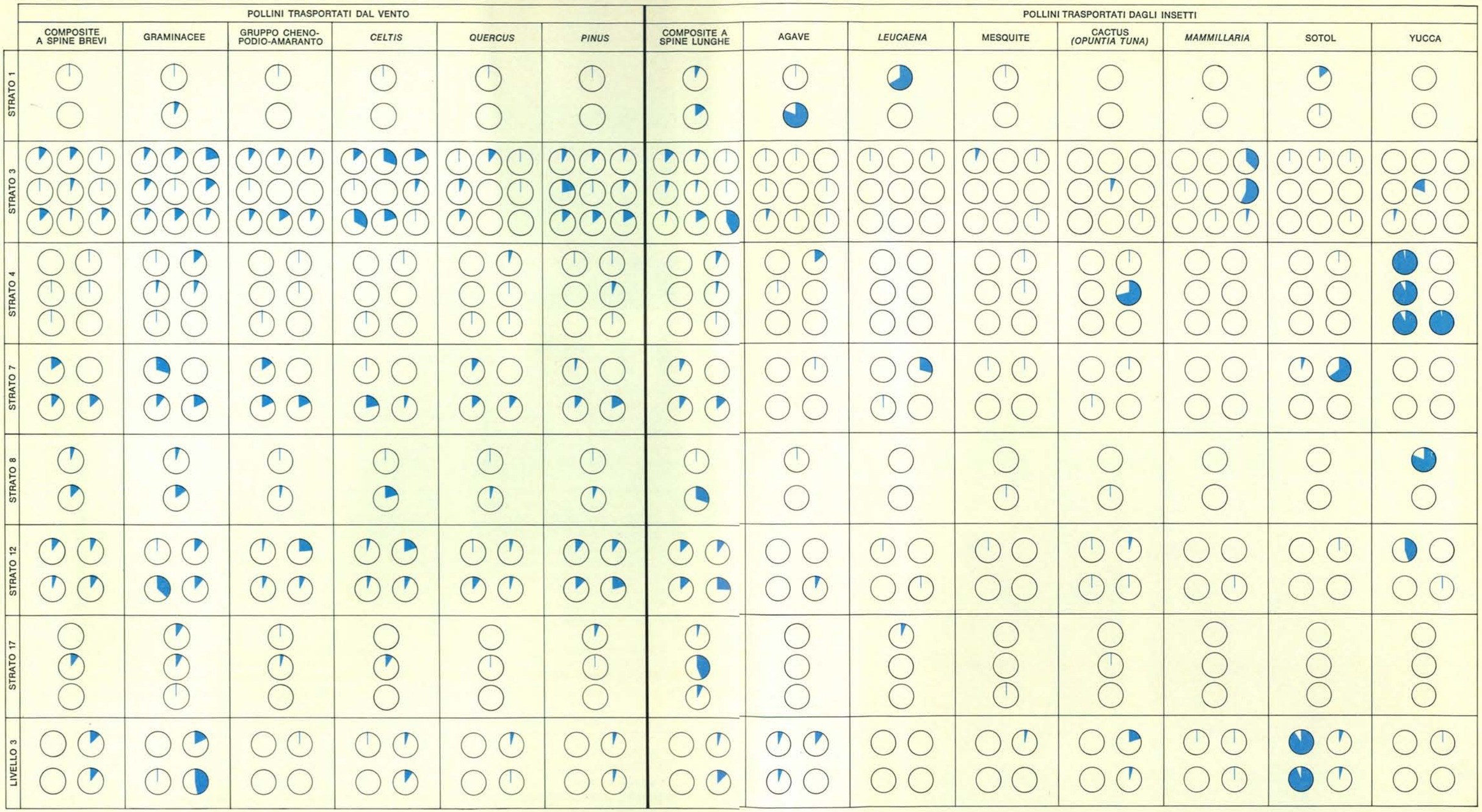
Occasionalmente, nel corso dell'analisi dei coproliti, si trovano prove di infestazione parassitaria. Callen e T. W. M. Cameron hanno dimostrato la presenza di uova di tenia in campioni risalenti a 3-5000 anni fa provenienti da una grotta in Perù, Huaca Prieta. In altri campioni contenuti in uno strato di 10 000 anni fa della grotta Danger nello Utah, Gary F. Fry della Youngstown State University e Edwin Englert jr. dell'Università dello Utah hanno trovato uova di un altro parassita intestinale, l'acantocefalo. I coproliti provenienti da Mesa Verde in Colorado contenvano uova di ossiuri. Per quanto riguarda zecche, acari, pidocchi e pulci che infestano il pelo del corpo umano, le testimonianze etnografiche riguardanti il comportamento di spidocchiamento dell'uomo bastano a rendere poco sorprendente il ritrovamento dei loro resti nei coproliti. Callen ha identificato zecche in campioni di uno strato di Tehuacan che risale a circa 6000 anni fa. Campioni provenienti da Mesa Verde e dalla Grotta Salata del Kentucky contengono acari e quelli della grotta dello Utah contengono pidocchi e lendini.

Soltanto negli ultimi decenni gli archeologi hanno cominciato a rendersi conto di quante informazioni si possano ottenere dall'analisi del polline. Come ben sanno i botanici, la membrana esterna o esina del granulo di polline possiede una notevole stabilità chimica, e pertanto i granuli di polline interrati sono praticamente indistruttibili. Inoltre la morfologia della membrana è determinata geneticamente, sicché la forma di un granulo di polline indica spesso da quale specie di piante esso proviene, o almeno da che genere o famiglia. A parte l'informazione che la comune analisi del polline può fornire all'archeologo, i granuli di polline, casualmente o deliberatamente ingeriti, attraversano il canale alimentare senza essere praticamente alterati. All'osservazione al microscopio i granuli di polline possono fornire altrettante informazioni sulla dieta preistorica quanto qualsiasi altro materiale vegetale. La specie che il polline rappresenta può indicare la stagione dell'anno in cui un sito era occupato e anche fornire un quadro generale dell'ambiente a quell'epoca.

Vediamo ora che cosa ci ha rivelato

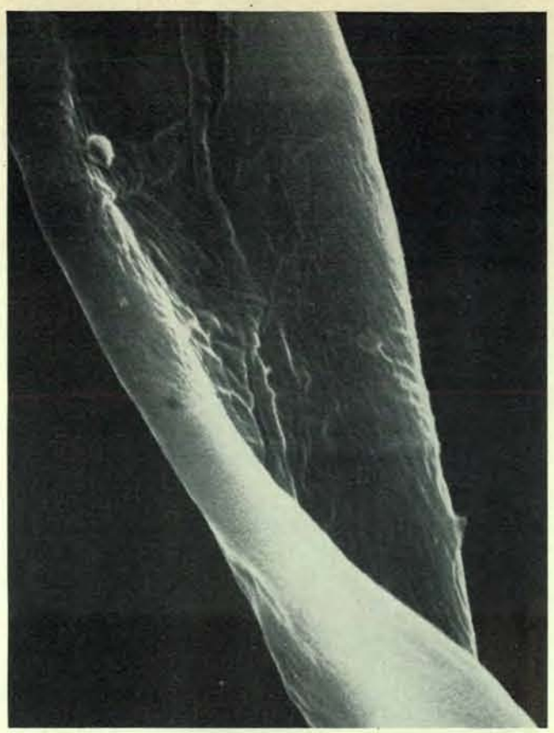
l'identificazione di polline e di altro materiale vegetale sulle attività degli uomini che vissero nei ripari di roccia del Texas precedentemente menzionati. I coproliti trovati in questo insediamento sono il risultato di un'occupazione umana intermittente che coprì un periodo di 1300 anni, da circa il 500 a.C. all'800 d.C. Dai vari strati di occupazione sono stati recuperati in totale 43 campioni: tutti tranne nove contenevano polline. Uno dei primi compiti di chi analizza il polline consiste nel separare i granuli in tre classi: quelli che sono trasportati dal vento, quelli che sono trasportati dagli animali (prevalentemente insetti) e quelli che provengono da piante che si autoimpollinano. (Alcune piante acquatiche liberano il loro polline sott'acqua, ma questi granuli in genere mancano di guscio e raramente si conservano; essi, quindi, non possono rientrare nel quadro di questa ricerca).

I granuli di polline trasportati dal vento vengono prodotti in grande quantità. Per fare un esempio, si ritiene che le foreste di abete della Svezia centrale e meridionale producano 75 000 tonnellate di polline all'anno. Un'antera di una pianta di questo tipo può produrre sino a 70 000 granuli di polline e raramente ne produce meno di 10 000. Nelle piante che vengono impollinate dagli insetti la produzione è inferiore, ed è di circa 1000 granuli per antera;

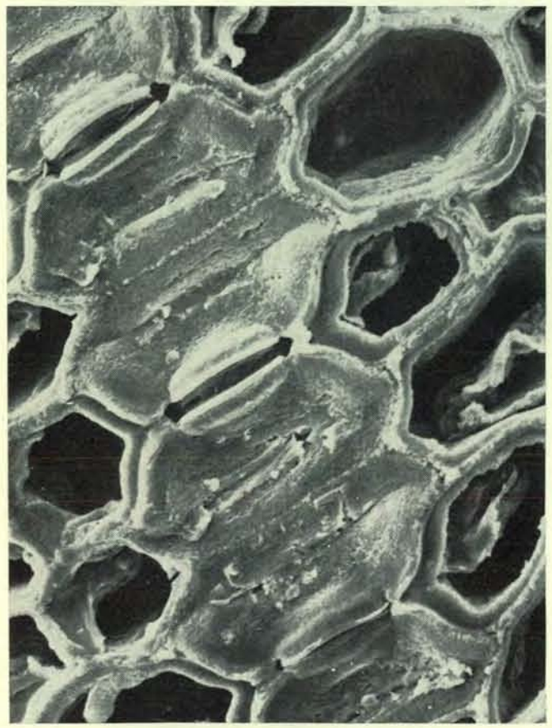


I principali pollini trovati in 34 dei coproliti portati alla luce in un riparo di roccia del Texas sudoccidentale provenivano da sei gruppi di piante che producono polline che viene trasportato dal vento (*a sinistra*) e da otto gruppi di piante che producono polline che viene trasportato dagli insetti (*a destra*). La presenza di grandi quantità di polline in parecchi campioni indica che gli indiani arcaici che visitavano periodicamente il sito ingerivano diverse specie di fiori succhiandole o masticandole per estrarne il nettare, facendone infuso o mangiandole in insalata. Soltanto in un campione un tipo di polline trasportato dal vento era presente insieme alla rispettiva pianta che era usata come cibo; l'ingestione di questi pollini era spesso casuale.

I semi rinvenuti nei coproliti possono far identificare la pianta da cui provengono. Questo seme di *quinoa* è ingrandito 70 volte.



Le fibre costituiscono un'altra classe di materiale vegetale molto significativa. Questa è una fibra di cotone, ingrandita 6000 volte.



Anche gli stomi sulla superficie delle foglie rendono possibile l'identificazione. Questa foglia di agave è ingrandita 800 volte.

Nella fotografia è ingrandito 3000 volte un cristallo di forma romboidale che si forma solamente nei tessuti d'*Agave lecheguilla*.

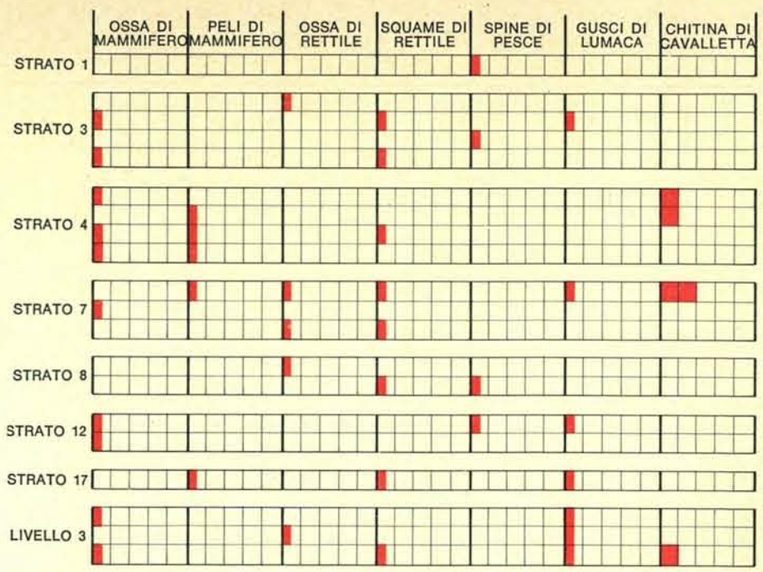
nelle piante che si autoimpollinano può essere di meno di 100 granuli per antera.

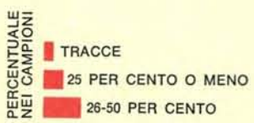
I sei principali pollini trasportati dal vento che si sono trovati nei campioni dei ripari di roccia del Texas appartenevano a tre generi e a tre più ampi gruppi di piante; la morfologia degli ultimi tre tipi di polline non ha permesso l'identificazione della specie o del genere che li aveva prodotti. A seconda della stagione, una « pioggia » più o meno regolare dei sei tipi di polline cadeva sul posto. Alcuni granuli furono probabilmente ingeriti perché erano caduti su sostanze alimentari; altri forse si accumularono nell'acqua da bere o anche furono inalati e poi inghiottiti.

I tre generi di piante identificate come fonti del polline sono *Celtis, Quercus* e *Pinus*. Sia le specie *C. laevigata* e *C. reticulata* sia la *Q. pungens* crescono ancora nelle terre alluvionali relativamente ben irrigate vicine all'insediamento. I più vicini pini conosciuti, tuttavia, sono gruppi di *Pinus cembroides* a circa 80 miglia a est.

Le piante di due su tre dei gruppi meno ben definiti che contribuirono alla pioggia di polline comprendevano alcuni membri della grande famiglia delle composite e membri della altrettanto grande famiglia delle graminacee. Il terzo gruppo è costituito da una costellazione di piante ben note ai botanici negli Stati Uniti sudoccidentali. Tale gruppo viene chiamato « Cheno-Am » perché i nomi ufficiali dei suoi due principali costituenti sono le chenopodiacee e il genere *Amaranthus*. I pollini prodotti dalle molte specie dei due gruppi di piante sono quasi impossibili da riconoscere l'uno dall'altro. Lo stesso si può dire per le molte specie di composite, se si eccettua il fatto che la membrana esterna (esina) dei granuli di polline di composite trasportati dal vento ha spine corte mentre l'esina dei pollini trasportati dagli animali ha spine lunghe.

Nei coproliti dei ripari di roccia la quantità di polline trasportato dal vento andava da zero sino al 48 per cento del totale. Tranne un caso, non ci sono prove che pollini trasportati dal vento fossero ingeriti in relazione alla dieta. L'eccezione è rappresentata dal campione che conteneva polline trasportato dal vento nella misura del 48 per cento. Il polline faceva parte del gruppo « Cheno-Am », e nel campione erano presenti anche molti semi di chenopodio. Questa pianta ha un lungo periodo di fioritura e perciò il polline degli ultimi fiori poteva facilmente mescolarsi con i primi semi che giungevano a maturazione quando questi ultimi





I resti animali nei coproliti provenienti dal riparo di roccia del Texas non sono molto abbondanti. Questo fatto, tuttavia, più che riflettere una carenza di proteine animali nella dieta degli indiani arcaici, è da attribuirsi probabilmente alla più efficace digestione della carne. Dal punto di vista quantitativo l'animale più diffuso nei reperti è la cavalletta; ma venivano anche mangiati piccoli roditori, lucertole e pesci.

venivano raccolti a scopo alimentare.

L'analisi dei pollini trasportati dal vento al riparo di roccia costituisce il fondamento di un'ipotesi sulle condizioni ambientali locali nel periodo che va dal 500 a.C. a circa l'800 d.C. Un certo quantitativo di tutti e sei i pollini trasportati dal vento è stato trovato, almeno in tracce, in uno o più campioni di tutti e otto gli strati del sito. Sebbene oggi il pino non sia presente nelle vicinanze del sito, è chiaro che, fatta questa eccezione, le piante di oggi sono le stesse del passato. Il polline in questione non ci dice nulla, tuttavia, sul numero relativo di piante di ciascun tipo durante tale intervallo di 1300 anni. È anche possibile che i pini non crescessero vicino al riparo di roccia e che il polline di pino ivi trovato fosse sospinto dal vento da pinete lontane.

Nei coproliti del riparo di roccia sono rappresentati anche otto generi o gruppi maggiori di piante che producono polline trasportato da insetti. Tra questi vi erano alcune composite caratterizzate da granuli di polline a spine lunghe, ma nell'insieme la flora era tipica degli ambienti aridi o semiaridi. Una specie riconoscibile era l'agave del deserto, *Agave lecheguilla*. Gli altri sei pollini definiscono solo i generi: *Leucaena, Prosopis* (mesquite), *Opuntia* (fico d'India), *Mammillaria, Dasylirion* (sotol), e un genere molto appariscente della famiglia delle liliacee, la yucca.

I fiori dell'agave, del sotol e della yucca venivano succhiati o masticati per estrarne il nettare, o se ne facevano infusi o venivano mangiati in insalata.

In genere un campione o più campioni, provenienti dai sette strati superiori lenticolari del riparo di roccia contenevano, per lo meno in tracce, polline di agave; nello strato 1 la percentuale di polline di agave in un campione raggiungeva l'81 per cento (*si veda la figura alle pagine 90 e 91*). Il polline di sotol è stato trovato in almeno un campione di tutti gli strati meno due. Nel deposito più antico del sito la percentuale di polline di sotol in due campioni superava il 90 per cento. Cinque degli otto strati hanno dato campioni che contenevano polline di yucca; negli strati 3 e 8 la percentuale superava l'80 per cento, e nello strato 4, in quattro dei sei campioni, era superiore al 90 per cento.

I fiori di cactus, a quanto pare, erano meno popolari. Il polline dei fiori di *Mammillaria* è stato trovato solo

in campioni provenienti da tre strati. Lo strato 3 conteneva la concentrazione massima: il 57 per cento. I fiori di fico d'India erano un po' più considerati: solo lo strato 1 era completamente privo del loro polline. Ciò nonostante, soltanto lo strato 4 (con un contenuto del 71 per cento nei campioni) e il deposito più antico (con un contenuto del 20 per cento) mostrano i segni di un consumo significativo di questo fiore. Quando si consideri che le due varietà di cactus crescono in habitat simili e fioriscono nella stessa stagione, è strano rilevare che solo quattro campioni contenevano entrambi i tipi di polline in tracce o più. Le scarse testimonianze di un consumo simultaneo dei due fiori possono essere espressione di una qualche sconosciuta preferenza alimentare o possono essere semplicemente legate al limitato numero di coproliti di cui disponiamo.

Oltre ai semi di chenopodio associati con il polline di tipo « Cheno-Am », la collezione conteneva molto altro materiale vegetale macroscopico. Gli abitanti del riparo di roccia non mangiavano solo i fiori, ma anche le foglie di molte piante monocotiledoni, come la yucca, il sotol e l'agave. Essi mangiavano anche i fiori di *Leucaena* e di mesquite (*Prosopis*). Come uno di noi ben sa per esperienza personale (Bryant) i fiori di mesquite sono molto amari se li si mangia crudi. Forse i fiori venivano bolliti per farne infuso o li si lasciava fermentare per renderli più gustosi. In alcuni campioni è stata trovata corteccia d'albero in quantità considerevole; quale trattamento potesse rendere accettabile la corteccia come alimento non sappiamo. Forse se ne faceva un uso medicinale.

La quantità di proteine animali ingerite dagli abitanti del riparo, se si eccettuano quelle contenute nelle cavallette, era forse piuttosto limitata (*si veda la figura a pagina 93*). Tra i materiali d'origine animale identificati nella nostra analisi vi sono ossa di pesciolini d'acqua dolce, scaglie di piccoli rettili (probabilmente lucertole) e conchiglie di lumaca.

Le testimonianze fornite dal materiale vegetale sulle stagioni in cui il riparo fu abitato comprendono alcune variabili. Nello strato 1, vicino alla superficie, lo scarso numero di coproliti fece sorgere in un primo momento dei dubbi sull'attendibilità del campione. Tuttavia, la presenza di polline di agave nella misura dell'81 per cento in uno dei campioni, e di polline di *Leucaena* nella misura del 66 per cento nell'altro, fa pensare che il riparo fosse occupato nella tarda primavera e nella prima estate. Tali piante fioriscono in questo periodo dell'anno. Un'indicazione analoga è fornita anche dalla quantità di polline di cactus e di yucca e di bulbi di cipolla selvatica nei due campioni. Il contenuto in polline trasportato dal vento conferma anch'esso questa ipotesi: sia il *Celtis*, sia il pino sono piante il cui polline si forma in primavera.

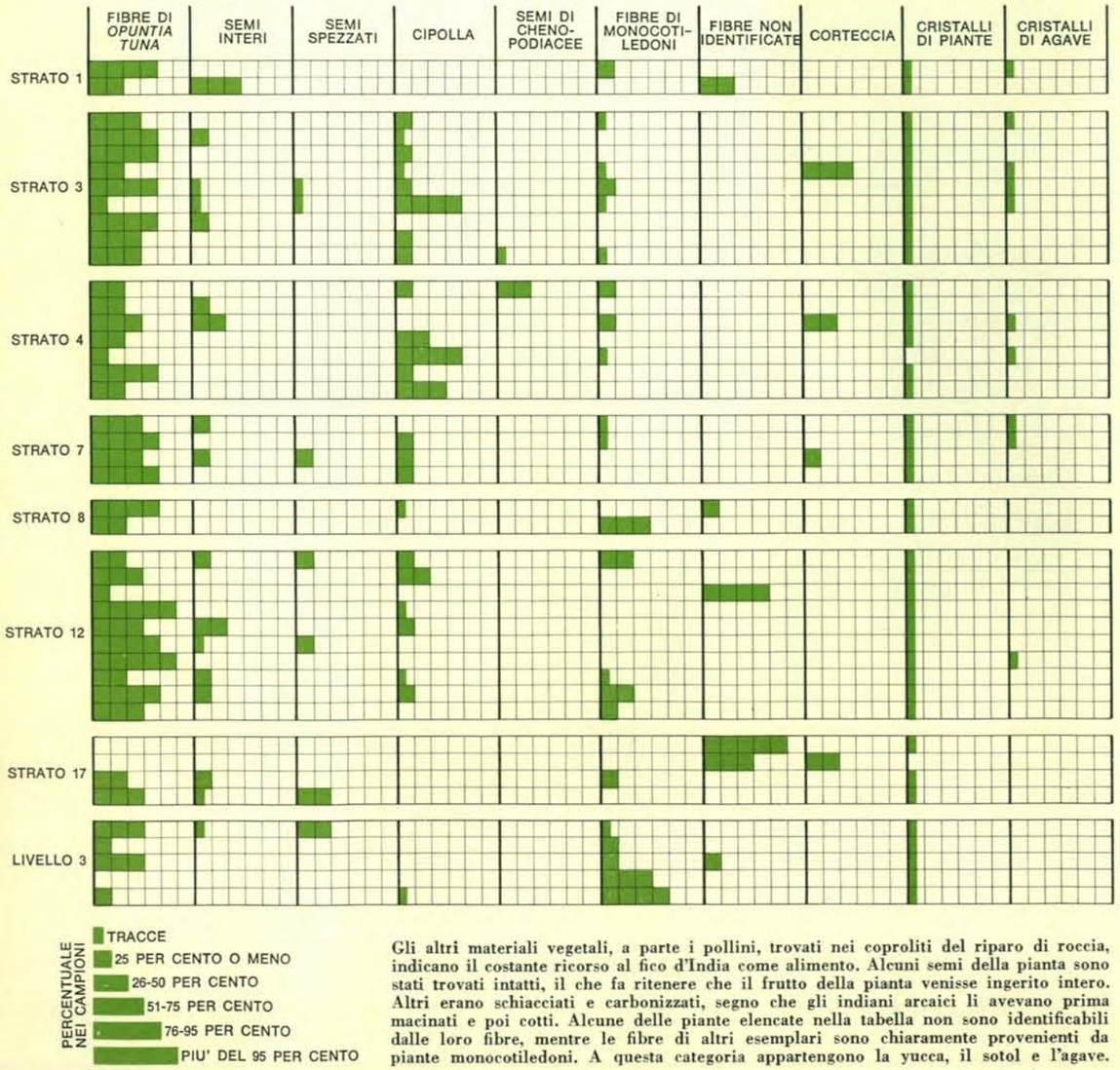
L'ipotesi di un'occupazione limitata alla tarda primavera e alla prima estate sembra valida anche per tutti gli altri strati dell'insediamento a eccezione di due. Lo strato 8 conteneva solo due campioni; poiché in uno di essi mancava qualsiasi materiale vegetale, il giudizio sulla stagione dell'occupazione è risultato impossibile. Nello strato 12 pur essendo assai scarsi i dati relativi ad altre varietà di polline, a parte quello trasportato dal vento, il rimanente materiale vegetale contenuto nel campione, come fibre di cactus e fibre di yucca, potrebbe essere stato raccolto e mangiato in qualsiasi periodo dell'anno. Poiché i frutti di cactus maturano nella tarda estate e poiché i semi di fico d'India sono presenti in vari campioni dello strato 12, si può ragionevolmente presumere che nel periodo in cui il deposito si formò il sito era occupato nella media e tarda estate. Nel complesso le testimonianze fornite dai coproliti sembrano indicare che il riparo di roccia del Texas veniva visitato ogni anno durante i mesi più caldi da una popolazione nomade di cacciatori e raccoglitori primitivi.

I più antichi coproliti trovati nel riparo risalgono a meno di 3000 anni fa. Che cosa potrebbero dirci dei campioni più antichi? Callen, prima di morire prematuramente, esaminò quattro campioni presumibilmente umani provenienti dal sito neanderthaliano di Lazaret in Francia. Il contesto della scoperta dei campioni di Lazaret fa ritenere che essi risalgano a 50 000-70 000 anni or sono. Nell'immergere i campioni Callen rilevò che nessuno di essi colorava la soluzione. In due egli poté identificare frammenti d'osso, peli e pezzi di carbone. Questa prova dell'uso del fuoco nella preparazione della carne lo convinse che almeno due di tali campioni erano con molta probabilità di origine umana. Egli non trovò in nessun campione materiale vegetale.

Recentemente, nel nostro laboratorio, abbiamo cominciato a lavorare su un vasto campionario di materiale molto più antico; si tratta di circa 500 coproliti presumibilmente umani provenienti da un sito della Francia mediterranea, Terra Amata. I campioni ci sono stati forniti dallo studioso Henry de Lumley (si veda l'articolo *Un accampamento paleolitico a Nizza*, di Henry de Lumley, in « Le Scienze », n. 13, settembre 1969). Se si eccettuano i coproliti, nessun resto umano è stato trovato in tale insediamento, ma gli strumenti litici ivi raccolti rivelano che gli occupanti di tale campo sul mare erano probabilmente rappresentanti dell'*Homo erectus*, il precursore dell'uomo moderno. De Lumley ritiene che gli strati portati alla luce a Terra Amata possano risalire a 300 000 anni fa. Mentre Callen aveva trovato che i più recenti campioni di Neanderthal non reagivano in immersione, alcuni dei campioni di Terra Amata, nel nostro laboratorio, hanno debolmente colorato la soluzione.

Molto resta ancora da fare con i reperti di De Lumley. Tuttavia, un'analisi preliminare ha identificato granuli di sabbia, che vengono quasi inevitabilmente ingeriti sulle spiagge, scaglie di carbone, che indicano l'uso del fuoco nella preparazione degli alimenti e frammenti di conchiglie di molluschi che mettono in evidenza una possibile fonte di risorse alimentari per gli abitanti. Finora non abbiamo trovato né ossa né residui vegetali, sebbene altre analisi condotte sui campioni di Terra Amata, intraprese anni fa nel laboratorio dell'Università di Aix-Marseille, abbiano dimostrato la presenza di diversi pollini trasportati dal vento e di polline della ginestra che è trasportato dagli insetti.

Oggi viene sempre più largamente riconosciuto quale contributo potenziale offrano le analisi di questo tipo all'archeologo che voglia accertare le preferenze alimentari in una varietà di contesti preistorici e che ricerchi un'indicazione relativamente precisa sulla stagione dell'anno in cui gli insediamenti transitori erano occupati. Il nostro laboratorio riceve oggi regolarmente campioni da analizzare che sono stati scoperti dai ricercatori in tutte le parti del mondo. Man mano che il lavoro procede, potremo forse dare altre indicazioni sul comportamento preistorico dell'uomo altrettanto inaspettate quanto la nostra scoperta che gli indiani arcaici del Texas sudoccidentale erano, per lo meno in senso alimentare, « figli dei fiori ».



Gli altri materiali vegetali, a parte i pollini, trovati nei coproliti del riparo di roccia, indicano il costante ricorso al fico d'India come alimento. Alcuni semi della pianta sono stati trovati intatti, il che fa ritenere che il frutto della pianta venisse ingerito intero. Altri erano schiacciati e carbonizzati, segno che gli indiani arcaici li avevano prima macinati e poi cotti. Alcune delle piante elencate nella tabella non sono identificabili dalle loro fibre, mentre le fibre di altri esemplari sono chiaramente provenienti da piante monocotiledoni. A questa categoria appartengono la yucca, il sotol e l'agave.

STUDI BOMPIANI

Angelo Baracca
Silvio Bergia
La spirale
delle alte energie

Tra i frutti essenziali della recente "rivoluzione culturale" studentesca c'è la consapevolezza della non neutralità della scienza. In questo libro, tale assunto si applica alla fisica e al suo nucleo "fondamentale", lo studio delle particelle, dove il meccanismo di sviluppo porta verso macchine sempre più gigantesche, esperimenti costosissimi, personale pletorico: verso "la spirale delle alte energie", tutta condizionata dalla strumentazione e dai rapporti produttivi ad esclusione di ogni altra alternativa possibile. Il testo, rigoroso ma sempre comprensibile, esplora la struttura della disciplina e analizza minuziosamente i meccanismi di sviluppo (e di condizionamento e di potere) che imprigionano la fisica delle particelle. L. 4.000

I problemi di cui la cultura tradizionale non si occupa. STUDI BOMPIANI

GIOCHI MATEMATICI

di Martin Gardner

Sui paradossi creati dalle relazioni non transitive

Quando di una relazione R possiamo dire che ogni volta che vale tra x e y e tra y e z vale anche tra x e z, cioè che da xRy e da yRz scende xRz, allora definiamo transitiva la relazione in questione. Per esempio la relazione « minore di » è una relazione transitiva sui numeri reali. Se 2 è minore di π e la radice quadrata di 3 è minore di 2, allora la radice quadrata di 3 è minore di π. Anche l'uguaglianza è una relazione transitiva: se $a = b$ e $b = c$, allora $a = c$. La vita quotidiana ci fornisce altri esempi di relazioni transitive, come « prima di », « più pesante di », « più alto di », « dentro a » e molte altre.

È facile incontrare relazioni che non sono transitive. Se A è il padre di B e B è il padre di C, non accadrà mai che A sia il padre di C. Se A ama B e B ama C, non è detto che A ami C. I giochi più familiari abbondano di regole transitive (nel poker, se la combinazione del giocatore A vince quella di B e la combinazione di B quella di C, allora la combinazione di A vince quella C), tuttavia alcuni giochi possiedono regole non transitive (o intransitive). Si prenda come esempio quel gioco diffuso tra i bambini in cui, dopo aver contato fino a tre, si presenta all'avversario la mano con le dita chiuse a pugno, per indicare la « roccia », oppure la mano con tutte le dita ripiegate tranne indice e medio, per indicare le « forbici », oppure tutte le dita stese per indicare la « carta ». Ora la roccia rompe le forbici,
carta e la carta avvolge la
questo gioco la relazione
intransitiva.

Accade talvolta in matematica, in modo particolare nella teoria delle probabilità
avere a che fare con una relazione che, contrariamente alle aspettative, non è transitiva. Se
mente in contrasto con l'intuizione da lasciarci perplessi, si verifica ciò che chiamiamo un paradosso non transitivo.

Di questi il più antico e il più noto è un paradosso connesso con la votazione, conosciuto anche come paradosso di Arrow dal nome di Kenneth J. Arrow, dato il suo ruolo cruciale nel « teorema di impossibilità » di Arrow, che valse a quest'ultimo il premio Nobel per l'economia
opera *Social Choice and Individual Values* (1951), Arrow specificò cinque condizioni che quasi tutti ammettono debbano essere soddisfatte da una democrazia in cui le decisioni sociali sono basate su preferenze individuali espresse col voto. Arrow dimostrò che queste cinque condizioni sono logicamente contraddittorie. Non è possibile concepire un sistema di voto che non violi in cer
ste cinque condizioni fondamentali. In conclusione, un sistema di voto perfettamente democratico non è possibile in linea di principio.

Secondo le parole di Paul A. Samuelson: « La ricerca di una democrazia perfetta svolta dai grandi pensatori della storia a noi nota si rivela ora come la ricerca di una chimera, per una intrinseca contraddittorietà logica... A questo punto, studiosi di tutto il mondo – matematici, politici, filosofi ed economisti – stanno cercando di salvare il salvabile dalla scoperta devastante di Arrow, che nei confronti della
teorema di incompletezza di Gödel del 1931 nei confronti della logica matematica.

Esaminiamo ora il paradosso della votazione considerando innanzitutto un difetto fondamentale
do attuale di elezione dei funzionari pubblici, per il quale si verifica spesso l'insediamento in una carica di un uomo che è cordialmente odiato dalla maggioranza dei votanti ma che possiede una minoranza di seguaci entusiastici. Supponiamo che il 40 per cento dei votanti siano seguaci entusiasti del candidato A. L'opposizione si divide in un 30 per cento di sostenitori di B e 30 per cento di sostenitori C. In questa situazione viene eletto A anche se il 60 per cento dei votanti sono contrari.

Si crede comunemente che per evitare queste conseguenze della suddivisione dei voti sia sufficiente permettere ai votanti di ordinare secondo le preferenze individuali l'insieme dei candidati. Sfortunatamente anche in questo caso si determinano decisioni irrazionali. La matrice presentata nella figura in basso a sinistra mostra il noto paradosso della votazione nella sua forma più semplice. La riga superiore indica che un terzo dei votanti preferisce, nell'ordine, i candidati A, B e C. La riga di mezzo indica che un altro terzo li preferisce nell'ordine B, C e A, mentre l'ultimo terzo nell'ordine C, A e B, come è indicato dall'ultima riga. Esaminate ora con attenzione la matrice e notate che, quando i candidati sono confrontati due a due, la non transitività fa nuovamente capolino. Due terzi dei votanti preferisce A a B, due terzi B a C e due terzi C ad A. Dunque A vince se è confrontato con B, B vince se è confrontato con C, C vince se è confrontato con A. Sostituite agli uomini delle proposte politiche e vedete come è facile per il partito al potere manipolare una votazione scegliendo semplicemente quale
in lizza debba essere messa ai voti per prima.

Il paradosso è noto in Francia come effetto Condorcet, dal nome del Marchese di Condorcet che per primo lo scoprì nel 1785. Lewis Carroll, autore fra l'altro di alcuni pamphlet sul problema del voto, lo riscoprì. La maggior parte dei primi sostenitori della rappresentanza proporzionale erano completamente ignari dell'esistenza di questo tallone d'Achille; in realtà gli esperti di scienze politiche non presero coscienza dell'esistenza del paradosso fino alla metà degli anni quaranta quando Duncan Black, un economista gallese, lo riscoprì in connessione al suo monumentale lavoro sui comitati dotati di potere decisionale. Attualmente gli esperti sono ben lontani dal poter stabi-
da Arrow vada reprimantemente
e in quella delle decisioni, di miglior sistema di voto. Una sorprendente soluzione raccomandata per situazioni di stallo da molti esperti di
la non transitività è tal
teoria delle decisioni
gliere a caso un « dittatore » che le
solva. Qualcosa del genere si verifica effettivamente in alcune democrazie, per esempio in Inghilterra, dove un monarca costituzionale (scelto a caso nel senso che l'ereditarietà non garantisce speciali disposizioni) ha la facoltà, attentamente delimitata, di risolvere certe situazioni senza sbocco che si verificano in circostanze eccezionali.

Il paradosso del voto può aver luogo in ogni situazione in cui si debba scegliere tra due alternative, dove le alternative variano su un insieme di tre o più. Supponiamo che A, B e C siano tre uomini che hanno chiesto contemporaneamente la mano di una ragazza. Le righe della matrice del paradosso del
strare come la ragazza classifichi ognuno degli uomini rispetto a ognuna delle tre caratteristiche che, a suo parere, sono le più importanti: l'intelligenza, la bellezza fisica e la ricchezza. Confrontandoli due a due la povera ragazza scopre che preferisce A a B, B a C e C ad A. È facile vedere come possano sorgere conflitti analoghi quando si tratti di scegliere un lavoro o di

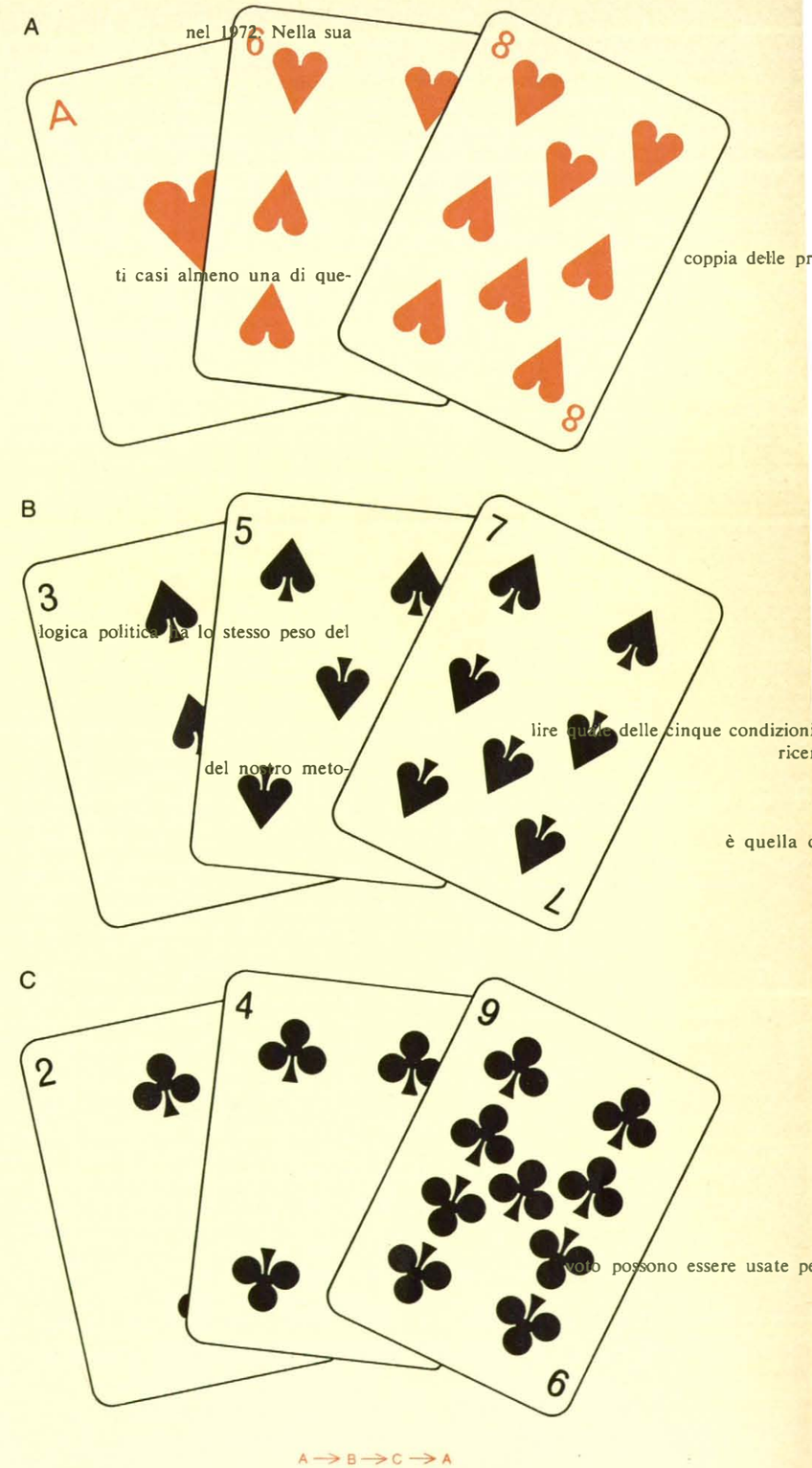
ORDINE DI PREFERENZA

VOTANTI	1	2	3
1/3	A	B	C
1/3	B	C	A
1/3	C	A	B

Il paradosso della votazione.

SQUADRE	D	E	F
A	8	1	6
B	3	5	7
C	4	9	2

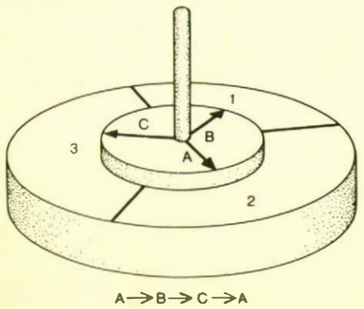
Il paradosso del torneo basato su un quadrato magico.



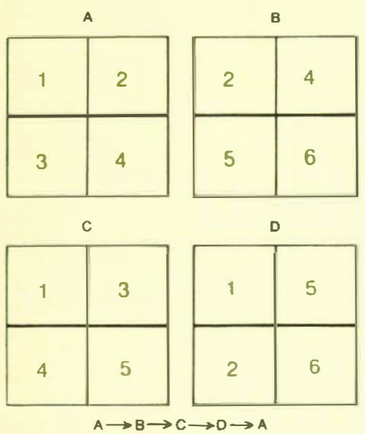
Gioco truccato basato su scommesse ricavato da un quadrato magico: A → B → C → A.

stabilire dove passare una vacanza.

Una deliziosa interpretazione della matrice è stata suggerita una volta da Paul R. Halmos. Supponiamo che *A*. *B* e *C* stiano rispettivamente per la torta di mele, la torta di mirtilli e la torta di ciliege. Un ristorante, per ogni pasto, ne offre solo due. Le righe mostrano come un avventore le classifichi rispetto a queste tre caratteristiche: gusto, freschezza e dimensioni della porzione. È perfettamente logico dal punto di vista dell'avventore, secondo Halmos, preferire la torta di mele a quella di mirtilli, la torta di mirtilli a quella di ciliege, la torta di ciliege a quella di mele.



La trottola non transitiva.



Carte per bingo non transitive.

B \ A	TT	TC	CT	CC
TT		1/2	1/4	1/2
TC	1/2		1/2	3/4
CT	3/4	1/2		1/2
CC	1/2	1/4	1/2	

Probabilità di vittoria di B.

Non vi è accordo fra gli esperti in merito alla frequenza con cui si presentino ordinamenti non transitivi come questi nella vita di tutti i giorni, ma alcuni recenti studi di psicologia e di economia indicano che sono più comuni di quanto si potrebbe pensare. Esistono perfino resoconti di esperimenti condotti sui ratti secondo i quali, sotto certe condizioni, le coppie di scelte compiute dai singoli ratti sarebbero non transitive. (Si veda: Warren S. McCulloch, *A Heterarchy of Values Determined by the Topology of Nervous Nets*, « Bulletin of Mathematical Biophysics », Volume 7, 1945, pag. 89--93.) Simili paradossi si verificano in quei tornei di tennis a squadre in cui ogni partecipante
incontra con ogni partecipante di ogni altra squadra. Supponiamo che nove tennisti siano classificati assegnando loro numeri dall'uno al nove, al migliore il numero nove e al peggiore il numero uno. La matrice mostrata nella figura di destra di pagina 96 è il noto quadrato magico di ordine tre. Le righe *A*, *B* e *C* indicano come i nove giocatori siano divisi in tre squadre, dove ogni riga rappresenta una squadra. Supponiamo che nei tornei a squadre, dove ogni membro di una squadra gioca con ogni membro delle altre, il giocatore più forte vinca sempre. Risulta che la squadra *A* batte la *B*, la *B* batte la *C* e la *C* batte la *A*, in ogni caso per cinque game contro quattro. È imposssibile dire quale squadra sia la migliore. La stessa non transitività si rileva nel caso in cui siano le colonne *D*, *E* e *F* a costituire le squadre.

Leo Moser e J. W. Moon hanno analizzato insieme molti di questi paradossi, su alcuni dei quali sono fondati certi divertenti e poco noti giochi basati su scommesse in cui le probabilità di vittoria sono tutte dalla parte di un giocatore, anche se ciò è tenuto nascosto dalla struttura del gioco. Per esempio, supponiamo che ogni riga (oppure ogni colonna) di una matrice quadrata costituita da un quadrato magico di ordine tre sia un insieme di carte da gioco, cioè sia *A* l'insieme costituito da asso, sei e otto di cuori, *B* da tre, cinque e sette di picche, *C* da due, quattro e nove di fiori (*si veda la figura della pagina precedente*). Ognuno di questi insiemi di carte è disposto a casaccio sul tavolo con il dorso verso l'alto. L'avversario, ignaro del funzionamento del gioco, viene invitato a tirare una carta scegliendola da un insieme qualsiasi, quindi tocca a voi pescare una carta da un altro insieme. La carta più alta dà la vittoria. È facile dimostrare che, da qualunque insieme di carte il vostro avversario peschi, potete sempre scegliere un insieme che vi dà la possibilità di vincere cinque volte su quattro. L'insieme *A* batte *B*, *B* batte *C* e *C* batte *A*. È possibile anche lasciare alla vittima la scelta se sia la carta più alta o quella più bassa a vincere. Se giocate avendo convenuto che sia la carta più bassa a vincere, è sufficiente che scegliate l'insieme vincente considerando un ordinamento ciclico non transitivo dei tre insiemi che giri nell'altro verso. È consigliabile giocare facendo uso di tre mazzi di carte diversi, in cui i dorsi delle carte siano differentemente colorati. Il mazzo di nove carte che serve per il gioco viene ogni volta mescolato e quindi suddiviso in tre insiemi, in base al colore del dorso. Naturalmente il trucco che sta alla base del gioco è nell'isomorfismo tra questa situazione e quella che genera il paradosso del torneo di tennis.

Vi sono molti altri semplici giochi basati su scommesse in cui prevale la non transitività. (Si veda, in questa rubrica, nel fascicolo di settembre del 1971, la descrizione di un insieme di dadi non transitivi.) In alcuni casi, come in quello della trottola progettata da Andrew Lenard (*si veda la figura in alto a sinistra*), la non transitività è facile da comprendere. La parte inferiore della trottola è fissa mentre il disco superiore può girare. Ognuno dei due giocatori sceglie una freccia differente, quindi si fa ruotare la trottola (in una direzione qualsiasi) e la persona la cui freccia indica il settore contrassegnato dal numero più alto vince. Abbiamo allora che *A* batte *B*, *B* batte *C* e *C* batte *A*, in ogni caso con una probabilità di due a uno.

Nell'insieme di quattro carte da bingo progettato da Donald E. Knuth (*si veda la seconda figura a sinistra*) la non transitività è nascosta con intelligenza. Due giocatori scelgono ognuno una carta da bingo. Si estraggono a caso numeri compresi tra uno e sei senza rimpiazzamento, come accade solitamente nel bingo. Se il numero estratto compare su una carta lo si segna con un fagiolo secco. Il primo giocatore che completa una riga orizzontale vince. Naturalmente in questo caso i numeri sono semplicemente dei simboli, infatti è possibile rimpiazzarli

con un insieme qualsiasi di sei simboli differenti. Lascio ai lettori la dimostrazione mediante il calcolo delle probabilità del fatto che la carta *A* batte la *B*, la *B* batte la *C*, la *C* batte la *D*, la *D* batte la *A*. Con tre giocatori il gioco è transitivo, ma le probabilità di vittoria delle quattro possibili triple riservano una sorpresa.

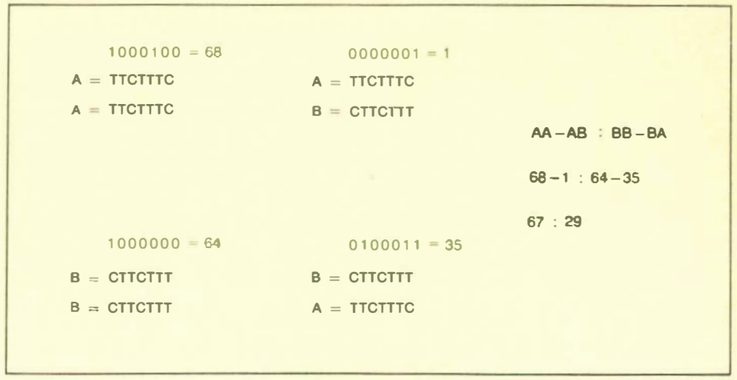
Una delle situazioni non transitive più assurde che si presentino nei giochi d'azzardo è stata scoperta dal matematico Walter Penney (nome alquanto appropriato) ed è stata presentata come problema nel « Journal of Recreational Mathematics » (ottobre, 1969, pagina 241). Si tratta di una situazione paradossale poco nota e in genere i matematici rimangono semplicemente sbalorditi quando ne vengono a conoscenza. È certamente uno dei più raffinati giochi d'azzardo truccati. Si può giocare utilizzando una moneta, una roulette puntando
nero oppure qualsiasi metodo che generi due alternative equiprobabili. Supponiamo di utilizzare una moneta e di tirarla tre volte. Vi sono otto possibili risultati equiprobabili: *TTT*, *TTC*, *TCT*, *TCC*, *CTT*, *CTC*, *CCT* e *CCC*. Ognuno dei due giocatori sceglie una tripla differente, quindi si lancia la moneta finché non appare una delle triple scelte stabilendo così la vittoria del giocatore corrispondente. Per esempio, se le triple scelte sono *TTC* e *CTC* e il risultato dei lanci è rappresentato dalla successione *CTTTC*, allora sulla base degli ultimi tre lanci si stabilisce la vittoria del giocatore che aveva scelto *TTC*. In conclusione, la vittoria spetta alla tripla che appare per prima tutta di seguito.

A questo punto si sarebbe disposti a credere che nessuna tripla ha maggiore probabilità di apparire prima di un'altra, ma basta una breve riflessione per rendersi conto del fatto che ciò non è vero nemmeno per le
Consideriamo le coppie *TT*, *TC*, *CT* e *CC*. *TT* e *TC* hanno uguali probabilità di presentarsi per prime perché, dopo che è apparso per la prima volta *T*, esso può essere seguito con la stessa probabilità da *T* come da *C*. Lo stesso ragionamento dimostra che *CC* e *CT* sono equiprobabili. Per motivi di simmetria abbiamo che *TT* = *CC* e *TC* = = *CT*. Tuttavia *CT* batte *TT* con probabilità di tre a uno e *TC* batte *CC* con uguali probabilità. Consideriamo *TC* e *CC*. La coppia *CC* viene sempre preceduta da *TC* tranne quando *CC* appare nei primi due lanci. Nel lungo periodo ciò accade solo una volta su quattro, quindi la probabilità che *TC* batta *CC* è di 3/4. Nella figura in basso della pagina a fronte sono confrontate a due a due tutte le possibili coppie e in ogni caso viene indicata la probabilità di vittoria di *B*, cioè del secondo giocatore.

B \ A	TTT	TTC	TCT	TCC	CTT	CTC	CCT	CCC
TTT		1/2	2/5	2/5	1/8	5/12	3/10	1/2
TTC	1/2		2/3	2/3	1/4	5/8	1/2	7/10
TCT	3/5	1/3		1/2	1/2	1/2	3/8	7/12
TCC	3/5	1/3	1/2		1/2	1/2	3/4	7/8
CTT	7/8	3/4	1/2	1/2		1/2	1/3	3/5
CTC	7/12	3/8	1/2	1/2	1/2		1/3	3/5
CCT	7/10	1/2	5/8	1/4	2/3	2/3		1/2
CCC	1/2	3/10	5/12	1/8	2/5	2/5	1/2	

Probabilità di vittoria di B *giocando con le triple.*

Le triple presentano una situazione ancora più sorprendente. Dato che non ha importanza che lato della moneta sia indicato come testa, sappiamo che *TTT* = *CCC*, *CCT* = *TTC*, *TCT* = = *CTC* e così via. Tuttavia, se esaminiamo le probabilità per coppie diverse, scopriamo che il gioco non è transitivo. Infatti, per ogni scelta di una tripla da parte del primo giocatore, il secondo può sempre sceglierne una migliore. La figura in alto di questa pagina fornisce, per ogni accoppiamento di triple, la probabilità che *B*, il secondo giocatore, vinca *A*. Per individuare la migliore risposta di *B* in corrispondenza a una tripla scelta da *A*, si determini



L'algoritmo di John Horton Conway per calcolare le probabilità di vittoria dello n-upla di B *su quella di* A.

la tripla di *A* prendendola nella prima riga orizzontale in alto, quindi si scenda lungo la colonna corrispondente a quella tripla finché non si incontri un numero in
babilità, quindi ci si sposti orizzontalmente verso sinistra fino alla prima colonna che riporta le triple preferite da *B*.

Si noti che la probabilità di vittoria di *B* è, al peggio, di 2/3 (ovvero di due a uno) e può arrivare a 7/8 (ovvero a sette a uno). Il caso di probabilità sette a uno è facile da capire. Si considerino *CTT* e *TTT*. Se *TTT* compare in una situazione che non sia l'inizio del gioco, allora deve essersi presentato in precedenza un *C*, il che vuol dire che *CTT* è stato già ottenuto. Ciò significa che *TTT* vince solo quando appare nei primi tre lanci. Ovviamente ciò accade solo una volta su otto tiri.

Barry Wolk, dell'Università di Manitoba, ha scoperto una regola originale per determinare la tripla migliore. Sia *X* la prima tripla scelta. Trasformiamola in un numero binario mutando *T* in 0 e *C* in 1. Dividiamo il numero ottenuto per 2 e arrotondiamo il quoziente al numero intero più vicino, moltiplichiamo quindi per 5 e aggiungiamo 4. Esprimiamo il risultato ottenuto nel sistema binario e ritrasformiamo le ultime tre cifre in *T* e in *C*.

La non transitività è conservata per tutte le *n*-uple con *n* maggiore di 3. Una tabella fornita da Wolk mostra le probabilità di vittoria di *B* per ogni possibile accoppiamento di quadruple (*si veda la figura di questa pagina*) Come nel caso delle due tabelle precedenti, e come accade anche per le *n*-uple di ordine più elevato, la matrice è simmetrica rispetto al centro. Il quadrante in alto a destra è quello in basso a sinistra ribaltato e lo stesso vale per quello in basso a destra e quello in alto a sinistra. Le probabilità della miglior risposta di *B* alla scelta di *A* sono indicate da un numero in colore.

B \ A	TTTT	TTTC	TTCT	TTCC	TCTT	TCTC	TCCT	TCCC	CTTT	CTTC	CTCT	CTCC	CCTT	CCTC	CCCT	CCCC
TTTT		1/2	2/5	2/5	3/10	5/12	4/11	4/11	1/16	3/8	3/8	3/8	1/4	3/8	7/22	1/2
TTTC	1/2		2/3	2/3	1/2	5/8	4/7	4/7	1/8	9/16	9/16	9/16	5/12	9/16	1/2	15/22
TTCT	3/5	1/3		1/2	3/5	5/7	1/2	1/2	5/12	5/12	9/16	9/16	5/14	1/2	7/16	5/8
TTCC	3/5	1/3	1/2		3/7	5/9	2/3	2/3	5/12	5/12	9/16	9/16	1/2	9/14	7/12	3/4
TCTT	7/10	1/2	2/5	4/7		1/2	1/2	1/2	7/12	7/12	5/14	1/2	7/16	7/16	7/16	5/8
TCTC	7/12	3/8	2/7	4/9	1/2		1/2	1/2	7/16	7/16	1/2	9/14	7/16	7/16	7/16	5/8
TCCT	7/11	3/7	1/2	1/3	1/2	1/2		1/2	1/2	1/2	9/16	5/12	7/12	7/12	7/16	5/8
TCCC	7/11	3/7	1/2	1/3	1/2	1/2	1/2		1/2	1/2	9/16	5/12	7/12	7/12	7/8	15/16
CTTT	15/16	7/8	7/12	7/12	5/12	9/16	1/2	1/2		1/2	1/2	1/2	1/3	1/2	3/7	7/11
CTTC	5/8	7/16	7/12	7/12	5/12	9/16	1/2	1/2	1/2		1/2	1/2	1/3	1/2	3/7	7/11
CTCT	5/8	7/16	7/16	7/16	9/14	1/2	7/16	7/16	1/2	1/2		1/2	4/9	2/7	3/8	7/12
CTCC	5/8	7/16	7/16	7/16	1/2	5/14	7/12	7/12	1/2	1/2	1/2		4/7	2/5	1/2	7/10
CCTT	3/4	7/12	9/14	1/2	9/16	9/16	5/12	5/12	2/3	2/3	5/9	3/7		1/2	1/3	3/5
CCTC	5/8	7/16	1/2	5/14	9/16	9/16	5/12	5/12	1/2	1/2	5/7	3/5	1/2		1/3	3/5
CCCT	15/22	1/2	9/16	5/12	9/16	9/16	9/16	1/8	4/7	4/7	5/8	1/2	2/3	2/3		1/2
CCCC	1/2	7/22	3/8	1/4	3/8	3/8	3/8	1/16	4/11	4/11	5/12	3/10	2/5	2/5	1/2	

Probabilità di vittoria di B giocando con le quadruple.

Studiando queste figure Wolk ha scoperto un'altra anomalia sorprendente quanto la non transitività che è connessa ai cosiddetti tempi d'attesa. Il tempo d'attesa per una certa
numero di lanci che in media occorre, nel lungo periodo, perché si presenti la *n*-upla in questione. Col passare del tempo nell'attesa dell'autobus diventa più breve il tempo d'attesa previsto. Tuttavia le monete non hanno memoria, quindi il tempo d'attesa per una *n*-upla è indipendente dai lanci precedenti. Il tempo d'attesa per *T* e *C* è 2. Per quanto riguarda le coppie, il tempo d'attesa per *TC* e *CT* è 4 ed è 6 per *TT* e *CC*. Per quanto riguarda le triple, il tempo d'attesa è 8 per *TTC, TCC, CTT, CCT*; 10 per *TCT, CTC*; 14 per *TTT, CCC*. Per ogni coppia formata a partire da queste triple, le probabilità di presentarsi per prima delle triple costituenti
conflitto coi tempi d'attesa. Tuttavia, nel caso delle quadruple, vi sono sei coppie che danno origine a contraddizioni. Per esempio, *CTCT* ha come tempo d'attesa 20 e *TCTT* ha tempo d'attesa 18. Tuttavia la probabilità che *CTCT* si presenti prima di *TCTT* è di 9/14, comunque superiore a 1/2. In altre parole, un evento che nel lungo periodo è meno frequente, ha maggiore probabilità di presentarsi prima di uno che è più frequente. In questo caso non vi è alcuna contraddizione logica, tuttavia si mostra che il «tempo d'attesa medio» gode di particolari proprietà.

Vi sono molti modi di calcolare le probabilità che una *n*-upla ha di precederne un'altra: si possono sommare serie infinite, tracciare diagrammi ad albero oppure utilizzare tecniche ricorsive che producono insiemi di equazioni lineari e così via. Una delle tecniche più strane ed efficienti è stata elaborata di recente da John Horton Conway della Università di Cambridge. Non ho idea del motivo per cui funzioni: semplicemente fornisce il risultato giusto, in modo quasi magico, cosa del resto comune a molti altri algoritmi inventati da Conway.

La chiave della procedura di Conway è il calcolo di quattro numeri binari che egli chiama numeri guida. Supponiamo che *A* sia la 7-upla *TTCTTTC* e *B* la 7-upla *CTTCTTT*. Vogliamo valutare la probabilità che ha *B* di precedere *A*. A questo scopo scriviamo *A* sopra *A*, *B* sopra *B*, *A* sopra *B* e *B* sopra *A* (*si veda la figura in basso di pagina 99*). Costruiamo ora, secondo le seguenti istruzioni, un numero binario che disponiamo sopra la 7-upla superiore di ogni coppia. Consideriamo la prima coppia, *AA*. Guardiamo la prima lettera della 7-upla superiore e chiediamoci se le sette lettere, iniziando dalla prima, corrispondono esattamente alle prime sette lettere della 7-upla inferiore. Ovviamente è così, quindi mettiamo un 1 sopra la prima lettera. Consideriamo ora la seconda lettera della 7-upla superiore e chiediamoci se le sei lettere rimanenti, iniziando da questa, corrispondono alle *prime sei* lettere della 7-upla inferiore. Ovviamente non è così, quindi poniamo 0 sulla seconda lettera. Le cinque lettere a partire dalla terza lettera della 7-upla superiore corrispondono alle prime cinque della 7-upla inferiore? No, quindi poniamo anche sulla terza lettera la cifra 0. Anche la quarta lettera è contrassegnata da 0. Giunti alla quinta lettera della 7-upla superiore, vediamo che le lettere restanti,
alle prime tre della 7-upla inferiore, quindi poniamo 1 sulla quinta lettera. La sesta e la settima lettera sono contrassegnate da 0. Il «numero guida *AA*», o semplicemente *AA*, è 1000100. In esso ogni 1 corrisponde a una risposta affermativa
tiva. Traducendo 1000100 nel sistema decimale otteniamo 68 come numero guida di *AA*.

La figura in basso della pagina precedente mostra il risultato di questa procedura quando si calcolino i numeri guida di *AA, BB, AB, BA*. Quando una *n*-upla è confrontata con se stessa, la prima cifra del numero guida deve essere ovviamente 1, quando è confrontata con un'altra *n*-upla può essere 1 o 0.

La probabilità che ha *B* di precedere *A* è data dal rapporto *AA—AB: BB——BA*. In questo caso abbiamo 68—1: 64—35=67: 29. Come esercizio il lettore può cercare di calcolare la probabilità che ha *CTT* di precedere *TTT*. I quattro numeri guida sono *AA*=7, *BB*=4, *AB*=0 e *BA*=3. Una volta inseriti questi numeri nella formula *AA——AB:BB—BA*, otteniamo una probabilità di 7--0:4--3, ovvero di 7:1, come previsto. L'algoritmo funziona altrettanto bene su *n*-uple di lunghezza diversa fra loro. Se applichiamo la nostra formula, per esempio, ad *A*=*TC* e a *B*=*T*, otteniamo 3—0:1—1=3:0, ovvero una probabilità di 1, che *T* batta *TC*.

Concludo con un problema di David L. Silverman, che è stato il primo a presentare il paradosso di Penney nella rubrica dedicata ai problemi che egli cura per il «Journal of Recreational Mathematics». I lettori non dovrebbero faticare a trovare la soluzione valendosi dell'algoritmo di Conway prima che sia pubblicata il mese prossimo. *CCTT* e *TTT* hanno come tempo d'attesa rispettivamente 16 e 14. Di queste due *n*-uple quale ha maggiore probabilità di presentarsi per prima e qual è il valore di questa probabilità?

Nel disegno rompicapo di Sam Loyd presentato il mese scorso compariva una torta quadrata di Washington da dividersi in sei quadrati, non necessariamente uguali. I 16 quadratini in cui la torta era suddivisa traevano in inganno dato che l'unica soluzione richiede una suddivisione in 9 quadrati, come una scacchiera per il ticktacktoe. Fatto ciò è facile dividere la torta in un quadrato di 2 per 2 e in cinque quadrati unitari.

Il problema della piramide magica del dottor Matrix, presentato nell'articolo di febbraio, ha affascinato molti lettori. Hiram Fuller Gutgasz ha sottolineato il fatto che la soluzione del problema è unica in un senso ancora più forte di quello che avevo supposto. Infatti non è necessario restringere il campo dei numeri assegnabili agli spigoli ai numeri positivi minori o uguali a 10, dato che la soluzione è unica per *ogni* insieme di otto numeri interi positivi distinti. Due lettori, J. A. McCallum e Sheldon B. Akers, si sono entrambi domandati se fosse possibile costruire una piramide magica sfruttando otto numeri interi positivi consecutivi. Entrambi sono giunti alla soluzione e hanno dimostrato la sua unicità. I numeri sono 4, 5, 6, 7, 8, 9, 10, 11 e la costante magica utilizzata 24.

Nell'articolo di aprile sui numeri esagonali e sui numeri stellari si chiedeva se un numero esagonale poteva essere un cubo. David Chess, Sin Hitotumatu e Wesley Johnston sono stati i primi, tra molti lettori, a mostrare come si potesse dare facilmente una risposta negativa a questo problema. Come ho mostrato nell'articolo, ogni numero esagonale può considerarsi come la differenza di due cubi consecutivi, quindi il problema si riduce a stabilire se l'equazione $(x+1)^3-x^3=y^3$ possiede una soluzione intera. Se riscriviamo l'equazione ponendo $x^3+y^3=(x+1)^3$, vediamo subito che si tratta di un caso particolare dell'ultimo teorema di Fermat. In questo caso, in cui gli esponenti sono uguali a 3, è stato dimostrato da molto tempo che non esiste una soluzione intera dell'equazione.